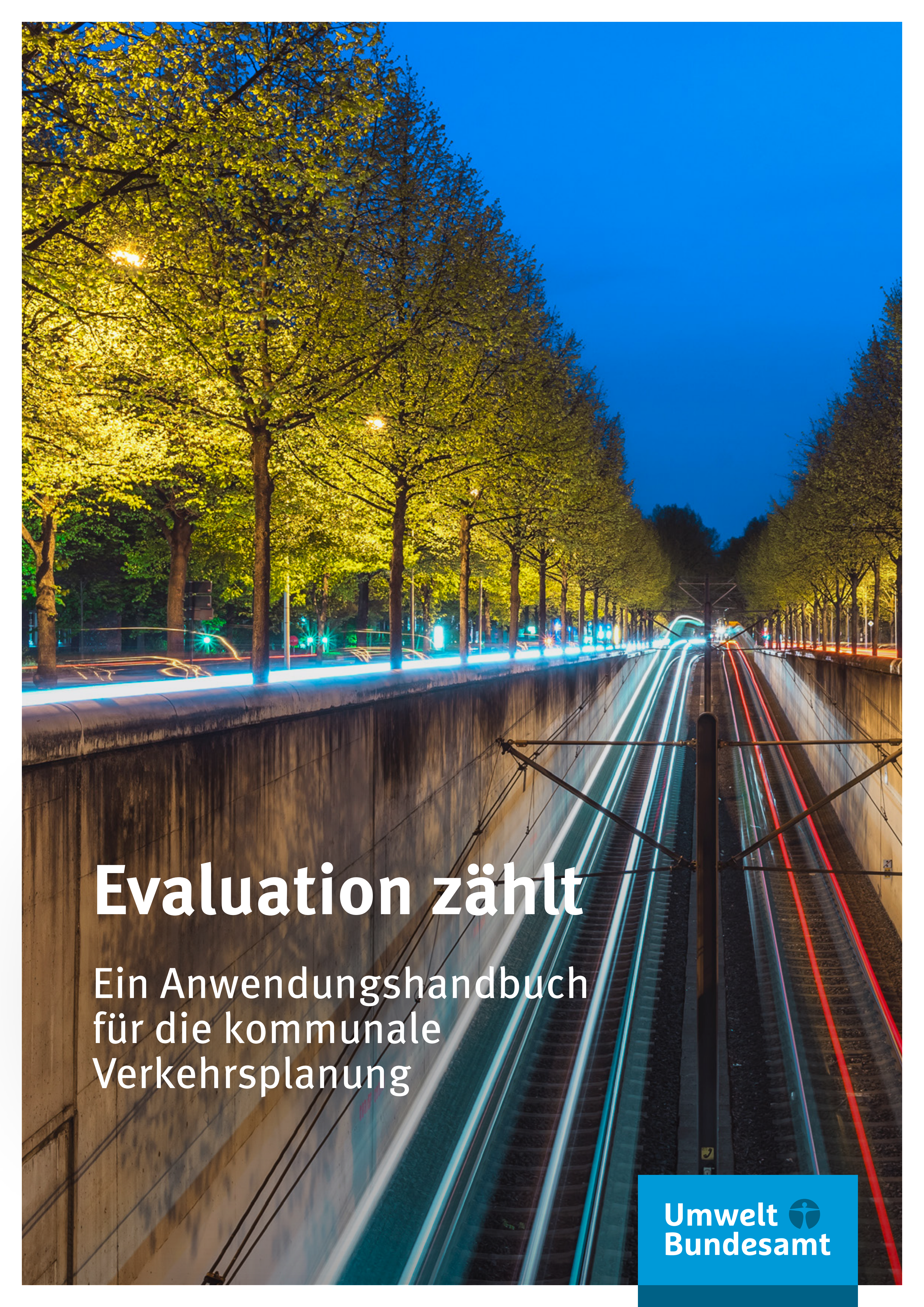


Evaluation zählt

Ein Anwendungshandbuch für die kommunale Verkehrsplanung

Umwelt Bundesamt

Herausgeber:
Umweltbundesamt | Fachgebiet I 3.1 Umwelt und Verkehr
Postfach 14 06 | 06844 Dessau-Roßlau
Wörlitzer Platz 1 | 06844 Dessau-Roßlau
Tel: +49 340-21 03-0
info@umweltbundesamt.de
www.umweltbundesamt.de


/umweltbundesamt.de
/umweltbundesamt

Autoren:
Katrin Dziekan
Veronique Riedel
Nicola Moczek
Stephan Daubitz
Stephanie Keßler
Stefanie Kettner
Michael Abraham

Direkter Kontakt zu den Autorinnen und Autoren:
katrin.dziekan@uba.de

Publikationen als pdf:
http://www.uba.de/evaluation-verkehrsplanung

Cartoons:
Auke Herrema | www.aukeherrema.nl

Übersetzung:
Vertaalbureau Snelvertaler | www.snelvertaler.nl
sowie Nicola Mozek | www.psyplan.de

Titelfoto: panoramarx – Fotolia

Satz:
Team VIERZIG A | www.vierzig-a.de

Stand: Januar 2015

ISSN 2363-832X

Evaluation zählt

Ein Anwendungshandbuch für die kommunale Verkehrsplanung



Inhaltsverzeichnis

Vorwort 4

1 Einführung 6

1.1 Wo bitte liegt Evalstadt? 6

1.2 Was ist eine Evaluation? 7

1.3 Was ist der Sinn einer Evaluation? 8

1.4 Wie läuft eine Evaluation ab? 9

Wann wird evaluiert?, S. 9 | Welche Arten von Evaluationen gibt es?, S. 11

1.5 Wer sollte bei der Evaluation eingebunden werden? 12

1.6 Weshalb gibt es dieses Buch? 12

2 Wirkungsevaluation 14

2.1 Maßnahmenziele 15

2.2 Ursache und Wirkung 18

2.3 Indikatoren 21

2.4 Untersuchungsdesigns 23

Quasi-experimentelle Vergleichsgruppen, S. 25 | Untersuchungsdesigns ohne Vergleichsgruppen, S. 26 | Entscheidung für ein Evaluationsdesign, S. 27

2.5 Methoden der Datenerhebung 29

Ausführliche Einzelbefragungen, S. 29 | Umfragen mit Fragebogen, S. 30 | Fokusgruppen, S. 32 | Beobachtungstechniken, S. 32 | Vor- und Nachteile der Methoden zur Datenerhebung, S. 33 | Skalen, S. 34 | Größe und Auswahl der Stichproben, S. 35 | Grundlegende Anregungen für die Durchführung einer Umfrage, S. 37

3 Dateneingabe und Datenauswertung 37

3.1 Datenaufbereitung und Datensicherung 37

3.2 Datenverarbeitung und -analyse 41

Nullhypothese und Alternativhypothese, S. 43 | Bestimmung des Signifikanzniveaus, S. 44 | Wahl des richtigen Testverfahrens für die vorliegenden Daten, S. 44 | Berechnung der Testparameter, S. 45 | Testentscheidung, S. 45 | Interpretation, S. 45

3.3 Qualitative Datenanalyse 45

Transkription und Exzerpt der aufgezeichneten Daten, S. 45 | Codierung der Texte, S. 46 | Vergleich und Verallgemeinerung, S. 47

3.4 Ergebnisdiskussion 47

4 Prozessevaluation 49

4.1 Ziele und Ergebnisse der Prozessevaluation 49

4.2 Widerstände und Treiber 50

4.3 Methoden der Datenerhebung 52

Standardisierte Formulare, S. 53 | Storytelling und Learning History, S. 54 | Sonstige Methoden , S. 58

5 Dokumentation und Bericht 59

5.1 Ergebnisdarstellung 59

5.2 Datendarstellung 60

Wie Daten nicht präsentiert werden sollten, S. 63

5.3 Übertragbarkeit 64

Beurteilung der Übertragbarkeit, S. 64 | Aufgaben der Ursprungsstadt, S. 65 | Die Aufgaben der Zielstadt, S. 66

6 Und jetzt: Zurück zur Realität 67
6.1 Einige Mythen über Evaluation 67
»Evaluationen sind viel zu kompliziert«, S. 67 | »Evaluationen sind zu teuer«, S. 67 |
»Evaluationen dienen nur dazu, das Projekt oder Programm zu bewerten und dann abzuschaffen oder die Verantwortlichen zu kritisieren.«, S. 68
6.2 Störungen im Ablauf einer Evaluation 68
Die Rahmenbedingungen für das Projekt verändern sich, S. 68
Das Projekt verzögert sich, S. 69 | Es werden zu viele Daten erhoben, S. 69 |
Es werden zu wenige Daten erhoben, S. 69
6.3 Die Qualität einer Evaluation erhöhen 69
Realistische Erwartungen formulieren, S. 69 | Realistische Abläufe planen, S. 69 |
Rückmeldungen geben und einholen, S. 69 | Dauerhafte Speicherung der Daten sicherstellen , S. 70

7 Das allerletzte Kapitel 72

8 Glossar 72

9 Literatur und Internetseiten 75
9.1 Kommentierte Literaturliste zu Evaluationen 75
9.2 Webseiten 76
9.3 Kommentierte Literaturliste zu Datenerhebung und Analyse 77
Software für Online-Erhebungen, S. 78 | Software für die Datenauswertung, S. 78 |
Software für qualitative Datenauswertung, S. 79
9.4 Datenschutz 79
9.5 Allgemeine Literatur 79

10 Vorhandenes Datenmaterial und Statistiken 80
10.1 Einwohner- und Beschäftigungsdaten 80
10.2 Umweltdaten 80
Luftschadstoffe, S. 80 | Lärm, S. 80 | Treibhausgas, S. 81
10.3 Kraftfahrzeugdaten 81
10.4 Verkehrsdaten 81
Straßenverkehrszählungen, Straßendatenbanken, Unfallstatistiken, S. 81 |
Statistiken für den öffentlichen Personenverkehr, S. 81
10.5 Verhaltensdaten 81
Mobilität in Deutschland (MiD), S. 81 | Mobilität in Städten SrV 2008 (2013), S. 82 |
Deutsches Mobilitätspanel, S. 82 | Kraftfahrzeugverkehr in Deutschland 2010, S. 82 |
Fahrleistungserhebung 2002 (2014), S. 82

11 Abbildungen und Referenzen 82

12 Die Autorinnen und Autoren 83

Vorwort

Was bringt eigentlich die Einführung eines Fahrradverleihsystems? Wie viele neue Fahrgäste für den öffentlichen Verkehr werden durch eine Marketingkampagne gewonnen? Funktionieren die stadtweiten Verkehrssicherheitsmaßnahmen in Schulumgebungen auch in anderen Städten?

Mit solchen und ähnlichen Fragestellungen zur Wirkungsabschätzung von verkehrlichen Maßnahmen sehen sich Akteure in der täglichen Arbeit konfrontiert. Im EU-Projekt CIVITAS MIMOSA standen die kommunalen Mitarbeitenden in den Städten Bologna, Funchal, Gdansk, Tallinn und Utrecht ebenfalls vor dieser Herausforderung. Evaluationen von komplexen Situationen sind aber nicht einfach. Daher erarbeitete das Projektevaluationsteam der Technischen Universität Berlin, Fachgebiet Integrierte Verkehrsplanung, gemeinsam mit dem internationalen CIVITAS POINTER-Team, das mit der horizontalen Aufgabe der Koordinierung aller Evaluierungen in den 25 teilnehmenden Städten in CIVITAS PLUS (2008–2013) betraut war, ein Handbuch mit dem Titel »Evaluation matters«. Dieses unterstützt seit dem Erscheinen 2013 zahlreiche internationale Projektteams bei der Planung und Durchführung von Evaluation kommunaler verkehrlicher Maßnahmen weit über das CIVITAS Programm der EU hinaus.

Das Umweltbundesamt möchte dieses Wissen jetzt auch der deutschsprachigen Leserschaft zur Verfügung stellen und hat daher eine Übertragung des Handbuches in Angriff genommen. Der Anspruch war es, neben deutschen Quellen für Daten und Kontakte auch gute, praxisnahe Beispiele zu finden, die für die anschauliche Darstellung der Evaluationsschritte geeignet sind. Leider mussten wir die intensive Suche in Bezug auf geeignete Beispiele nach einiger Zeit erfolglos einstellen. Zum Einen gibt es auch in Deutschland bislang zu selten Evaluationen und zum Anderen waren die gefundenen Beispiele oft nicht auf andere – vor allem auch mittlere und kleinere Kommunen – übertragbar. Daher haben wir als Autorenteam eine fiktive Beispielstadt namens Evalstadt erschaffen und auf diese Art »perfekte« Evaluationsbeispiele erzeugt.

Wir hoffen, dass dieses Handbuch Ihnen das nötige Handwerkszeug liefert, um gute Evaluationen innerhalb Ihrer kommunalen Maßnahmen durchführen zu können. Bei einer zukünftigen Überarbeitung dieses Werkes greifen wir dann gerne auch auf Ihre methodisch fundierten und aussagekräftigen Beispiele zurück.
Gutes Gelingen!

Katrin Dziekan für die Autorinnen und Autoren


ICH DENKE..
WIR HABEN EIN PROBLEM


WIR WOLLEN, DASS UNSERE STADT EIN SCHÖNER ORT WIRD...
FÜR DIE MENSCHEN


MIT SAUBERER LUFT UND BÄUMEN
PLATZ FÜR RADFAHRER UND FUßGÄNGER


WAS STEHT ZWISCHEN UNS UND UNSEREM ZIEL?
ZU VIELE AUTOS!
© 2014 www.aukeherrema.nl


WAS SOLLEN WIR ALSO TUN?
MEHR ÖFFENTLICHE VERKEHRSMITTEL!


SPÄTER
IST ES DAS, WAS WIR WOLLEN
ICH GLAUBE, ETWAS IST SCHIEF GEGANGEN
FORTSETZUNG FOLGT

1 Einführung

Die meisten europäischen Städte setzen sich aktiv für eine nachhaltige Entwicklung ein. Insbesondere Maßnahmen rund um die Mobilität haben einen hohen Stellenwert bei der Entwicklung lebenswerterer Städte. In den vergangenen zehn Jahren fanden in diesem Sektor gewaltige Veränderungsprozesse statt. In fast allen Metropolen haben sich die Rahmenbedingungen für den so genannten Langsamverkehr deutlich verbessert. Zu Fuß gehen und Rad fahren ist auch in großen Städten wie Paris, London und Berlin nicht nur möglich, sondern kann Spaß machen. Car-Sharing, innovativer öffentlicher Personennahverkehr und E-Mobility erweitern das Spektrum der Fortbewegungsmöglichkeiten. Begegnungszonen oder so genannte shared spaces machen den Menschen neue Angebote für gemeinsam genutzte Flächen. Mobilität ist vielfältiger geworden – in den Metropolen, wie auch in vielen kleinen und größeren Städten. Aber wurde und wird das Ziel der Nachhaltigkeit erreicht? Sind die Städte und Kommunen auf dem richtigen Weg?

Diese Frage kann nur mit Hilfe von Evaluationen beantwortet werden. Das Handbuch zeigt daher konkrete Wege auf, wie auch Kommunen mit kleinem Budget gründliche Evaluationen planen und durchführen können. Es soll dazu beitragen, Beteiligte, Rahmenbedingungen und Einflussfaktoren kennen zu lernen, Planungsprozesse zu optimieren, Schwerpunkte zu setzen und fundierte Entscheidungen zu treffen – damit die umgesetzten Maßnahmen möglichst erfolgreich sind und die erreichten Ziele denen entsprechen, die ursprünglich erwartet wurden. Nicht zuletzt helfen gute Evaluationen auch dabei, aus Erfahrungen zu lernen.

Das Handbuch erklärt Schritt für Schritt das Vorgehen der Evaluation anhand von nachvollziehbaren Beispielen. Wo immer es sinnvoll erscheint, liefert es Hinweise zu weiterführenden Informationsquellen.

1.1 Wo bitte liegt Evalstadt?

Zur Verdeutlichung der Evaluation mit Beispielen wurde für diese Veröffentlichung »Evalstadt« erfunden. Es ist eine mittlere Stadt mit rund 90.000 Einwohnern irgendwo in Deutschland. Diese erfundene Evalstadt hat sich in besonderer Weise dem Thema Mobilität angenommen und unter dem Leitbild »Evalstadt – nachhaltig mobil« einen umfassenden Maßnahmenkatalog beschlossen. Ziel ist vor allem, den Pkw-Verkehr zugunsten nachhaltiger Verkehrsträger zu reduzieren. In dem Modellversuch »Mobile Evalstadt« geht es insbesondere um die Frage, ob sich verkehrs- bzw. mobilitätsbezogene Einstellungen und Verhaltensweisen durch verschiedene Strategien bzw. Maßnahmen verändern lassen. Dahinter steht die Annahme, dass sich Veränderungen des Verkehrsverhaltens erreichen lassen, wenn Veränderungen der Infrastruktur umgesetzt und neue Angebote geschaffen werden. Dabei sollen die Ziele und Maßnahmen begleitend kommuniziert werden damit die Bevölkerung überzeugt und motiviert wird. Für Evalstadt wurden 40 Aktivitäten konzipiert und durchgeführt. Das Spektrum der Maßnahmen reicht von kleineren Infrastrukturmaßnahmen, verbesserten ÖPNV-Angeboten und Informationskampagnen bis hin zu Maßnahmen für den Bereich Schule. Die Aktivitäten lassen sich folgenden Themenfeldern zuordnen:

- Alternativen zum Pkw (neue ÖPNV-Angebote/Dienstleistungen z. B. Erhöhung der Taktdichte oder bedarfsgerecht angepasste Linienführung, Förderung des Fuß- und Radverkehrs)
- Mobilitätsmanagement von Aktivitäten und Zielen (Mobilitätsberatung, ÖPNV-Angebote für Ausflüge und Freizeitaktivitäten, Aktionstage »Einkaufen ohne Auto«)
- Reduktion der Pkw-Nutzung (Ausbau von Carsharing-Angeboten und Verbreitung von Informationen über diese, von Medien begleitete Selbstversuche »Ein Monat ohne Auto«)
- Bewusstes Autofahren (z. B. Kurse für energiesparendes Autofahren)
- Mobilitätsbildung (für den Bereich Schule werden Unterrichtseinheiten »Nachhaltige Mobilität« konzipiert und an mehreren Schulen durchgeführt)

Natürlich werden in unserer fiktiven Evalstadt die implementierten Maßnahmen evaluiert. Wie dabei vorgegangen wurde zeigen wir in Auszügen in diesem Buch.

1.2 Was ist eine Evaluation?

Evaluationen – manchmal auch Evaluierungen genannt – haben das Ziel, Projekte oder Maßnahmen sach- und fachgerecht hinsichtlich vorher festgelegter Kriterien zu bewerten. Dazu werden systematisch relevante Informationen gesammelt und miteinander in Beziehung gesetzt. Informationen können in diesem Fall objektive Messergebnisse sein, z. B. die Konzentration von Luftschadstoffen oder ein Lärmpegel, Daten und Fakten wie der Preis eines Parkscheins oder die Anzahl an Parkplätzen, Statistiken wie Unfallhäufungen oder in Bussen beförderte Passagiere an bestimmten Tagen. Relevante Informationen können aber auch subjektive Bewertungen wie der Komfort eines Sitzes in der S-Bahn, die Zufriedenheit mit einer Beratung oder die wahrgenommene Sicherheit auf den Gehwegen vor einer Schule sein.

Die Entscheidung darüber, welches ein relevantes oder angemessenes Beurteilungskriterium ist, hängt von der Maßnahme und dem damit verbundenen Ziel ab. Es ergibt beispielsweise wenig Sinn, die Geschwindigkeit einer innerörtlichen Straßenbahn zu messen, wenn das Maßnahmenziel die Erhöhung der Auslastung ist. Falls es aber darum geht, den Lärmpegel zu senken oder die Taktfrequenz deutlich zu erhöhen, kann die Geschwindigkeit wichtige Hinweise geben. Diese Information alleine wird aber nicht ausreichen, um zu beurteilen, ob die Ziele aus diesem Beispiel erreicht wurden. Es müssen weitere sinnvolle Indikatoren oder Kriterien gefunden werden. In aller Regel werden daher mehrere Kriterien zur Bewertung einer Maßnahme herangezogen.

Eine Evaluation findet nicht nur zu einem Zeitpunkt statt, sondern ist im Idealfall ein kontinuierlicher Prozess. Sie startet bereits in der ersten Planungsphase und begleitet dann fortlaufend die Durchführung bis zum Ende der Maßnahme, ggf. sogar darüber hinaus.

Ein Beispiel aus Evalstadt: Zur Reduktion der PKW-Nutzung wird ein von Medien begleiteter Selbstversuch »Ein Monat ohne Auto« initiiert. Bewohnerinnen und Bewohner von Evalstadt lassen einen Monat lang freiwillig ihr Auto stehen und nutzen stattdessen nur alternative Fortbewegungsmittel. Der Selbstversuch soll mit Erfahrungsberichten der Teilnehmenden dieser Aktion in der Lokalzeitung und im Lokalradio dokumentiert werden. Als ein Erfolgskriterium wird die Anzahl der Teilnehmenden an dieser Maßnahme definiert, es sollen mindestens 30 sein. Die Kampagnenorganisation bewirbt die Aktion in den lokalen Medien mit Plakaten und sie schreibt lokale Vereine an. In diesem Beispiel wird nun die Wirksamkeit der Werbemaßnahmen auf die Zahl der Aktionsteilnehmenden überprüft. Acht Wochen vor dem Start wird die Zahl der Interessenten kontrolliert. Sind dann erst zwölf Interessierte und acht verbindliche Anmeldungen für die Teilnahme an der Kampagne zu verzeichnen, müssen weitere Maßnahmen zur Bewerbung unternommen werden, z. B. persönliche Ansprachen im Wohnviertel, Fußgängerzone etc. Ab sofort kann und muss die Kontrolle der Anmeldezahlen in viel kürzeren Abständen durchgeführt werden, immer mit dem Ziel, die 30 zu erreichen.

Das Beispiel soll verdeutlichen, dass Evaluation nicht mit Prüfung, Controlling oder Monitoring verwechselt werden darf. Diese Begriffe dürfen nicht als Evaluation missverstanden werden. Die Unterschiede sind subtil, aber dennoch wichtig. Prüfung oder Controlling meint die Überprüfung der Einhaltung von zwingenden gesetzlichen Vorschriften und Normen sowie die Budgetkontrolle. Somit ist sie ein Werkzeug zur internen Kontrolle. Das Monitoring wird normalerweise gleichzeitig mit der Umsetzung ausgeführt und dient der Überprüfung dieses Verfahrens, insbesondere hinsichtlich des Erreichens der angenommenen Arbeitsleistung und der Ergebnisse der ergriffenen Maßnahmen, sowie der für die Umsetzung eingesetzten Mittel.

Prüfung und Monitoring können als Informationsquelle genutzt werden. Sie werden damit Teil der Evaluation.

Zurück zum Beispiel »Ein Monat ohne Auto«. Der Verlauf der Aktion wird dokumentiert und hinsichtlich der Einhaltung der Regeln überwacht. Zur Prüfung, dass die Autos wirklich nicht benutzt werden, parken die Teilnehmenden diese auf einem abgesicherten Gelände. Während des Selbstversuchs führen die Teilnehmenden ein Mobilitätstagebuch.

1.3 Was ist der Sinn einer Evaluation?

Durch Evaluationen ist es möglich, die Wirkungen und Effekte der umgesetzten Maßnahmen quantitativ und qualitativ zu bestimmen. Sind die gewünschten Effekte eingetreten, dient die Evaluation zur Bestätigung der Vorgehensweise. Treten sie nicht oder nur teilweise ein, deckt die Evaluation Schwachstellen auf und kann dazu dienen, ergänzende Maßnahmen zu planen und umzusetzen. Evaluationen dienen also:

- der Bestimmung der Effekte einer Maßnahme,
- dem Lernen für künftige Projekte sowie
- dem Erfahrungsaustausch.

Das Organisationsteam stellt vier Wochen vor dem Start der Aktion »Ein Monat ohne Auto« fest, dass die Zahl der Angemeldeten auf 20 gestiegen ist. Bevor sie weitere Werbemaßnahmen planen und umsetzen, evaluieren sie die Zusammensetzung der Teilnehmenden. Dabei kommt heraus, dass diese nicht aus allen Stadtteilen von Evalstadt kommen. Insbesondere der Bezirk mit der höchsten Autonutzung ist bislang noch nicht vertreten. In der Folge werden die Werbestrategien auf diesen Bezirk zugeschnitten. Die Herangehensweise erweist sich als erfolgreich und am Beginn des Selbstversuchs starten 34 Freiwillige.

Wenn die Ergebnisse der Evaluation veröffentlicht werden, können auch andere von den Erfahrungen profitieren (siehe hierzu auch Kapitel 5.3 auf Seite 64 Übertragbarkeit). Funktioniert eine Maßnahme zur Lärmreduktion innerhalb Evalstadt gut, kann diese auch in eine andere Stadt übertragen werden. Liegen ausführliche Evaluationsergebnisse vor, können die Entscheidenden vor Ort direkt auf diese positiven Erfahrungen aus Evalstadt verweisen. Die Evaluation verhilft zu:

- einem besseren Verständnis für öffentliche Ausgaben,
- einer besseren Orientierung der Maßnahmenbündel auf bestimme Zielgruppen,
- einer Verbesserung zukünftiger Planungen und einer Optimierung bei der Zuweisung von Ressourcen.

Eine zielgerichtete Evaluation dient also sowohl den Entscheidungsträgern als auch der Gesellschaft.

KASTEN 1: Warum Misserfolg auch wichtig ist

Um ein definiertes Ziel zu erreichen, ist ein gewisses Maß an Arbeit erforderlich. Erfolge machen bewusst, was bereits geschafft wurde. Sie können dazu beitragen, dass wir zuversichtlicher und mit gesteigertem Engagement an (neue) Aufgaben herangehen. Wir neigen aber auch dazu, uns nach Erfolgen »auszuruhen«. Und nicht immer bedeutet ein (Teil-)Erfolg, dass tatsächlich alles gut ist.

Misserfolge und bislang unerreichte Ziele können auch motivierende Effekte haben. Sie erhöhen die Bereitschaft sich einzusetzen, vor allem dann, wenn die Erreichung des Ziels eine persönliche Wichtigkeit hat und wenn Misserfolge auf beeinflussbare Aspekte einer Situation zurückgeführt werden. Nur wenn sie als Zeichen für mangelnde eigene Fähigkeiten gewertet werden, kann dies die Motivation hemmen und Gefühle von Unzufriedenheit verstärken.

In jedem Fall kann die bewusste Auseinandersetzung mit unerreichten Zielen und mit Misserfolgen einen Lernprozess anstoßen.

Nach: Scholl, A. (2011). Motiviert durch Erfolg und Misserfolg. wissens.blitz



Eine Evaluation sollte niemals mit dem Ziel durchgeführt werden, Menschen zu kontrollieren oder Sanktionen vorzubereiten (z. B. die Rückzahlung von Fördergeldern). Ein solcher Schwerpunkt würde die grundlegenden Ziele der Evaluation gefährden und letztendlich die Einführung innovativer Maßnahmen blockieren.

1.4 Wie läuft eine Evaluation ab?

Wann wird evaluiert?

Idealerweise wird eine Evaluation kontinuierlich den gesamten Umsetzungsprozess begleiten, läuft also parallel zu Planung, Umsetzung und Betrieb der Maßnahme. Die nachfolgende Abbildung zeigt einen Überblick über die möglichen Phasen und die wichtigsten Tätigkeiten.

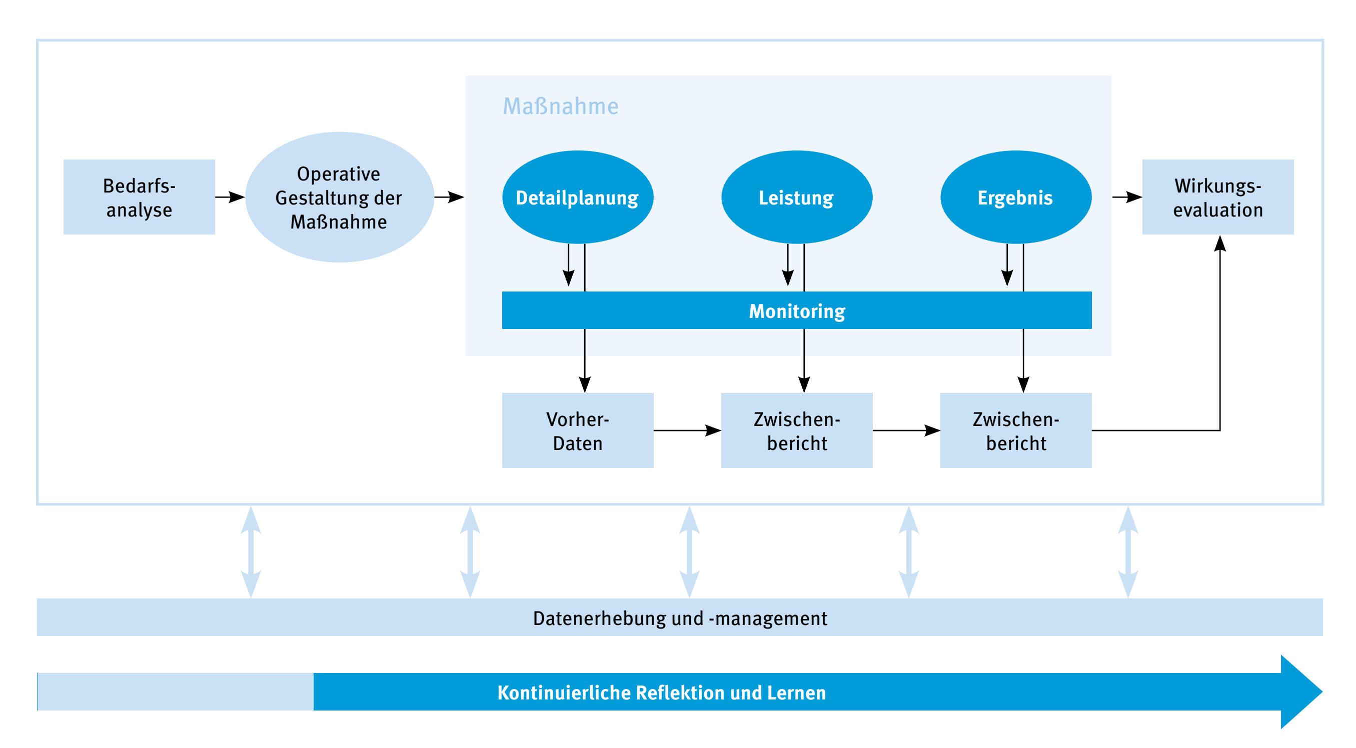


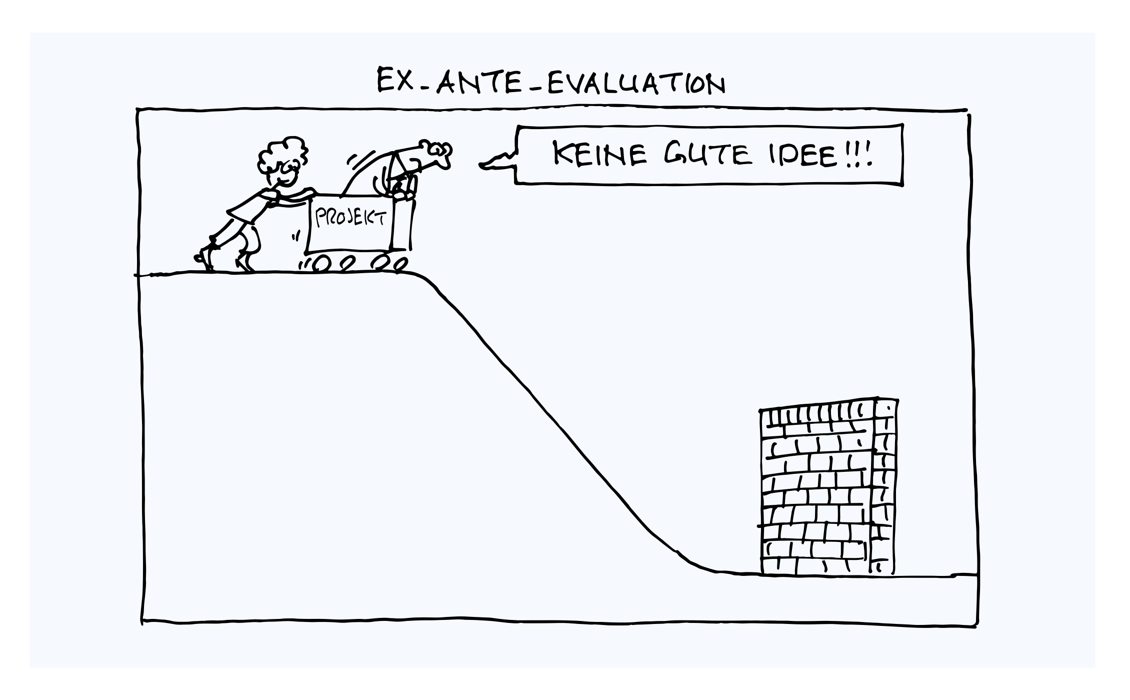
Abbildung 1-1: Phasen der Evaluation

Zu Beginn eines Planungsprozesses steht eine Bedarfsanalyse. Diese baut auf den identifizierten Problemfeldern auf, die häufig bereits aus einem bestehenden Leitbild abgeleitet wurden oder durch die (Verkehrs-/Umwelt-)Politik vorgegeben sind. In diesem Schritt wird der Bedarf beschrieben und überlegt, was zur Verbesserung der Situation getan werden kann. Dies führt dann zu einer Auswahl von Maßnahmen, die geeignet sein könnten, den Bedarf zu decken bzw. das gewünschte Ziel zu erreichen. Häufig gibt es mehr als eine Maßnahme zur Auswahl, aber die finanziellen Mittel sind beschränkt und es muss abgewogen werden.

In Evalstadt soll an einem Straßenabschnitt der Verkehrslärm reduziert werden. Dies könnte unter anderem durch eine Straßensperrung, die Errichtung von Lärmschutzwänden oder einer Reduzierung der Geschwindigkeit erreicht werden. Evalstadt möchte aber den Autoverkehr grundsätzlich reduzieren, indem die Attraktivität anderer Verkehrsmittel gesteigert wird. Daher wird an besagtem Straßenabschnitt eine Fahrspur je Richtung in einen Radschutzstreifen umgewidmet.

Diese Überlegungen sind Teil der Bedarfsanalyse, an deren Ende eine Vorbewertung der möglichen Maßnahmen steht. Dabei wird geprüft, was eine Maßnahme leisten kann. Es werden Annahmen über die Auswirkungen bestimmter Maßnahmen sowie zur Höhe der Kosten getroffen. Basierend auf den Ergebnissen wird entschieden, welche Maßnahme umgesetzt werden soll, um das Ziel zu erreichen. Diese Vorbewertung, die oft auch als »Ex-ante Evaluation« bezeichnet wird, kann dabei helfen, wirksame Entscheidungen zwischen unterschiedlichen Optionen zu treffen. Aber dies ist keine Evaluation in dem Sinne, wie wir sie in diesem Buch beschreiben, es handelt sich hierbei um eine Vorhersage oder eine Simulation von dem, was voraussichtlich geschehen wird. Mit einer Evaluation wird dokumentiert, was tatsächlich passiert.

In diesem Buch wird davon ausgegangen, dass dieser Schritt der Bedarfsanalyse bereits abgeschlossen ist. Im Folgenden wird daher vor allem die Wirkungsevaluation vorgestellt.



Nach der Ex-ante-Evaluation beginnt die operative Gestaltung der Maßnahme. Zu diesem Zeitpunkt wird auch die Evaluation geplant. Dazu gehören auch die Bestimmung des Zwecks der Evaluation, der verfügbaren Ressourcen und die Festlegung eines entsprechenden Evaluationsdesigns. Dies beinhaltet in jedem Fall die Bestimmung der Indikatoren und der Datengrundlage, an der die Veränderungen durch die Maßnahme gemessen werden können. Die Qualität der Datengrundlage legt dabei auch die Qualität der Datenauswertung fest. Bevor die Maßnahme implementiert wird, sollte auch schon eine

Vorerhebung durchgeführt werden. Diese Daten dienen dazu, die Situation vor der Einführung zu beschreiben und sind Grundlage für den späteren Vergleich.

Anschließend kann die Maßnahme gestartet werden, deren Fortschritt wird meist von einem Monitoring begleitet. Dies ist wichtig, um die laufende Umsetzung der Maßnahme zu verfolgen und zu bewerten. Häufig sind die dafür benötigten Daten auch für die Evaluation hilfreich oder umgekehrt. Die abschließende Wirkungsevaluation erfolgt nach Beendigung der Maßnahme, um zu beurteilen, wie gut die beabsichtigten Ziele erreicht wurden. Schließlich wird der Arbeitszyklus durch die Veröffentlichung der Ergebnisse und Erkenntnisse abgeschlossen.

Gerade bei längerfristigen Maßnahmen ist das Bewerten und Lernen über den ganzen Maßnahmenzyklus notwendig. Daher findet die Evaluation nicht einmal oder zweimal statt, sondern ist regelmäßiger Bestandteil bei der Umsetzung der Maßnahme.

Eine Evaluation benötigt Ressourcen: Zeit, Geld und beteiligte Personen. Aber sie ist immer ein sehr lohnender Prozess.

Welche Arten von Evaluationen gibt es?

Von großen, langfristigen, international vergleichenden Prozessen, die Millionen Euro kosten bis zu kleinen, kurzen Evaluationen einzelner Maßnahmen in einer Stadt gibt es eine große Vielfalt. Evaluationen können auf viele verschiedene Arten eingeteilt werden, für den Kontext in Bezug auf urbane Mobilität gibt es grundsätzlich zwei Bewertungsfelder – die Wirkungsevaluation und die Prozessevaluation, die immer parallel betrachtet werden.

Hauptziel bei der *Wirkungsevaluation* ist es, zu zeigen, ob es einen Effekt der Maßnahme gab bzw. ob die gesetzten Ziele erreicht wurden. Dazu wird die Situation vor der Maßnahme mit der Situation nach bzw. mit der Maßnahme verglichen. Die Wirkungsevaluation macht eine Bestandsaufnahme oder bewertet die Leistung. Üblicherweise benötigt man diese Art der Evaluation für die Entscheidungsfindung z. B. für die Auswahl oder den Umfang einer Maßnahme, da dabei »harte Fakten« präsentiert werden. Zu den alternativen Entscheidungen können beispielsweise gehören: den Eingriff an anderen Standorten durchführen (Übertragbarkeit), weiterhin Mittel bereitzustellen, die Mittel aufzustocken, den Status Probezeit beizubehalten, Änderungen durchzuführen und es erneut zu versuchen oder auch das Projekt zu stoppen. Die Wirkungsevaluation wird in Kapitel 2 dargestellt.



Die *Prozessevaluation* konzentriert sich auf die Mittel und die Verfahren bei der Umsetzung einer Maßnahme. Sie beginnt bei der Projektentwicklung und geht über die ganze Laufzeit des Projekts. Dabei werden für alle Projekttätigkeiten die positiven und negativen Faktoren, die den Umsetzungspro-

zess der Maßnahme beeinflussen, bewertet und so Informationen zur Überwachung und Verbesserung des Projekts geliefert. Die Prozessevaluation wird in Kapitel 3 vorgestellt.

Die Prozessevaluation und die Wirkungsevaluation sollen immer als Einheit gesehen werden. Nur ein sogenannter »gemischter Evaluationsansatz« mit beiden Verfahren kann eine echte Begründung für den Erfolg geben oder den Lernerfolg aus einem »Misserfolg« einer Maßnahme sicher stellen.

1.5 Wer sollte bei der Evaluation eingebunden werden?

An einer Maßnahme sind in der Regel viele Menschen mit unterschiedlichen Hintergründen, Interessen und Motivationen beteiligt. Eine gute Evaluation stellt daher sicher, dass möglichst alle Personengruppen, die von der Maßnahme möglicherweise betroffen sind, angemessen beteiligt werden. Hierzu zählen u.a.:

Politik und Entscheidungsträger – Das sind Mitarbeiterinnen und Mitarbeiter der Verwaltungen auf verschiedenen räumlichen und hierarchischen Ebenen, beispielsweise die Europäische Kommission, Ministerien und Behörden auf nationaler, regionaler und lokaler Ebene, usw. Für sie ist die Evaluation die Informationsquelle über das Projekt, seine Vorbereitung, Umsetzung und seine Ergebnisse.

Verwalter der Maßnahme – Das ist die Gruppe von Experten und Expertinnen, zu deren Aufgaben die Verwaltung der verschiedensten Aspekte des Projekts gehören. Durch die Ergebnisse der Evaluation erhalten sie Informationen über die Auswirkungen ihrer Arbeit, über positive Einflüsse und Schwierigkeiten.

Umsetzer – Das sind die Menschen, die die Maßnahme zu einem tatsächlichen Erfolg machen können. Sie können die Auswirkungen ihrer Arbeit in einem größeren Zusammenhang sehen und feststellen, wie sie zur allgemeinen Verbesserung der Lebensqualität in ihrer Stadt beiträgt.

Zielgruppen der Maßnahme – Das sind die potenziellen Nutznießer oder »Träger« einer Maßnahme. Diese Gruppe kann durch die Ergebnisse der Evaluation sehen, was sie vom Projekt erwarten können (im Vorfeld), und was im Rahmen des Projekts getan wurde (nach der Umsetzung).

Sonstige Interessengruppen – Das sind alle anderen, die nicht Teil der vorstehend genannten Gruppen sind. Auch ihnen sollten die Ergebnisse der Evaluation zur Verfügung gestellt werden. Transparenz verbessert die Akzeptanz einer Maßnahme.

Fachpersonal, d. h. Personen mit technischen und/oder methodischen Kenntnissen, sowie alle Personen, die bei der Definition der Fragen zur Evaluation oder bei der Interpretation der Evaluationsergebnisse helfen können. Die Beteiligung von unabhängigen Fachleuten kann für die Bereitstellung nützlicher Informationen für das Evaluationsteam und beim Ableiten von verallgemeinerbaren Erkenntnissen aus der Evaluationsstudie sehr hilfreich sein.

1.6 Weshalb gibt es dieses Buch?

Der Verkehr in Europa ist für nahezu ein Viertel der aktuellen energiebezogenen Treibhausgas-Emissionen verantwortlich, wobei Autofahrten mehr als drei Viertel aller gefahrenen Fahrzeugkilometer ausmachen. Viele Städte planen daher Maßnahmen, die beispielsweise zu einer verminderten Kfz-Nutzung führen und dadurch die negativen Auswirkungen des Verkehrs verringern sollen.

Der Anlass für das englische Buch »Evaluation matters«, was die Grundlage für das vorliegende Buch darstellte, war die Feststellung, dass europaweit die Anzahl und Qualität von Evaluationen erhöht werden könnte. Insgesamt gibt es bislang nur wenige Evaluationen im Bereich Mobilität. Zusätzlich werden die Ergebnisse häufig so veröffentlicht, dass lediglich Wissenschaftlerinnen und Wissenschaftler darauf zugreifen können, sie kommen also in der Praxis nicht in dem Maße an, wie es wünschenswert wäre.

Bei vielen Berichten fällt weiter auf, dass die Wirksamkeit der überprüften Maßnahmen eher schwach ausfällt. Dafür gibt es verschiedene Gründe:

- Manchmal wurden die Evaluationsmethoden nicht so gewählt, dass sie zufriedenstellende Ergebnisse liefern können. Hier fehlt es an fachlichem Know-how.

- In einigen Fällen scheint eine Evaluation von den Verantwortlichen nicht gewünscht, da sie mit einer Überprüfung oder Monitoring verwechselt wird. Als Resultat werden häufig Schwierigkeiten gar nicht oder nur unzureichend beschrieben. Das Potenzial zum Lernen aus diesen Evaluationen ist also sehr gering.
- In anderen Fällen wird die Evaluation grundsätzlich gewünscht, aber begrenzte Ressourcen können zu Ergebnissen mit geringem Wert führen oder die Evaluation behindern.

Insgesamt ist auffällig, dass in vielen Fällen der Nutzen, aber auch die Komplexität einer Evaluation deutlich unterschätzt wird. Dieses Buch soll dabei helfen, genau das zu ändern.



Das Buch entstand auf der Grundlage der Erfahrungen in den beiden Forschungsprojekten CIVITAS MIMOSA und CIVITAS POINTER, in deren Rahmen die Europäische Kommission integrierte Strategien für einen nachhaltigen städtischen Verkehr in den Städten Bologna, Funchal, Gdansk, Tallinn und Utrecht, sowie die Evaluierung aller Maßnahmen in den beteiligten Städte in CIVITAS PLUS (2008–2013) gefördert hat. In beiden Projekten wurde deutlich, dass es europaweit nicht genug Evaluationsliteratur gibt, die speziell auf den Kontext der urbanen Mobilität zugeschnitten ist. Die vielen positiven Reaktionen auf die englische Fassung, welche 2013 erschien, bestätigten diese Einschätzung.

Was dieses Buch kann:	nicht kann:
Es zeigt, dass Evaluation eine sinnvolle Aufgabe ist	Es kann nicht projektspezifische Richtlinien ersetzen
Es zeigt, dass die Evaluation von jedem gelernt werden kann	Bietet keine Leitschnur für eine komplette Programmevaluation
Es illustriert mit Beispielen alle Phasen des Evaluationsprozesses	Es gibt keine ausreichende Grundlage zur Evaluierung komplexer Infrastrukturprojekte
Es legt den Schwerpunkt auf kleinere kommunale Projekte	Es kann nicht das eigene Denken ersetzen

Tabelle 1-1: Was dieses Buch kann

Dieses Handbuch ist nicht darauf ausgelegt, Seite für Seite gelesen zu werden. Es ist vielmehr ein Leitfaden, der bei Bedarf als Referenz verwendet werden kann. Daher werden wichtige Punkte in verschiedenen Abschnitten wiederholt. Es arbeitet mit Fallstudien, Checklisten und Verweisen zu anderen methodischen Ansätzen, Online-Dokumenten sowie weiterführenden Informationsquellen.

Eine gute Evaluation ist eine komplexe Aufgabe, die aus Sicht der Autorinnen und Autoren aber allen möglich ist, die bereit sind, ein wenig Zeit und Anstrengung zu investieren und sich in einen offenen Dialog mit allen Beteiligten zu begeben – der Mehrwert lohnt sich.

2 Wirkungsevaluation

Mit Hilfe einer Wirkungsevaluation werden die messbaren Auswirkungen, also die Effekte einer Maßnahme bewertet. Eine Maßnahme kann dabei auch eine Intervention, ein Projekt oder eine Kommunikationskampagne darstellen. Typische Fragestellungen im Bereich von Mobilität und Verkehr sind:

- Steigt der Anteil der Radfahrenden auf einer bestimmten Strecke nach dem Einrichten von Radwegen?
- Verringert sich die Lärmbelastung an einer Straße durch die Verringerung der maximal zulässigen Geschwindigkeit von 50 km/h auf 30 km/h?
- Erhöht sich die wahrgenommene Sicherheit von Fußgängern, wenn der Zebrastreifen in der Nacht beleuchtet wird?

In aller Regel werden zur Bestimmung der Wirkung Daten erhoben. Um zuverlässig abschätzen zu können, ob die Wirkung tatsächlich auf die Maßnahme zurückzuführen ist, sollten relevante Informationen (also Daten) vor und nach der Maßnahme erfasst und miteinander verglichen werden. Idealerweise werden die Daten auch an einem Kontrollstandort erhoben um die Maßnahmenwirkung besser darzustellen.

Designs für Wirkungsevaluationen beruhen stets auf Vergleichen. In der Regel werden dazu zwei Gruppen verglichen, eine bei der die geplante Intervention stattfindet (Zielgruppe) und eine zweite, bei der keine Intervention vorgenommen wird (Kontroll- bzw. Vergleichsgruppe). Von Kontrollgruppe spricht man nur dann, wenn diese per Zufallszuteilung (auch Randomisierung) vor Beginn einer Maßnahme gebildet wird. Von Vergleichsgruppe spricht man hingegen dann, wenn die Zuteilung ohne Zufallszuteilung gebildet wird. Bei sonst gleichen Bedingungen für die beiden Gruppen sollten die gemessenen Ergebnisse nur der umgesetzten Maßnahme zugeschrieben werden. Eine Betrachtung der Zielgruppe zu ausschließlich einem einzelnen Zeitpunkt, z. B. nach der Maßnahme, ermöglicht keinerlei Aussagen über Wirkungen oder Veränderungen aufgrund der Maßnahme.

In Evalstadt wird ein Programm umgesetzt, das sich auf den schulischen Bring- und Holverkehr konzentriert. Ziel ist es, dass weniger Eltern ihre Kinder mit dem Auto zur Schule bringen. Wie kann die Wirkung der Maßnahmen gemessen werden?
Zunächst sollte überprüft werden, wie der Modal Split der Schulwege an den an der Maßnahme teilnehmenden Schulen nach einzelnen Jahrgangsstufen aussieht. Erfahrungsgemäß ist der Anteil der Autofahrten bei den Schulwegen in den höheren Klassen niedriger. Die Maßnahme wird anschließend in einer Klasse durchgeführt, jedoch in einer weiteren Klasse der gleichen Jahrgangsstufe, die als Kontrollgruppe dient, nicht. Nach Ablauf der Maßnahme wird der Modal Split erneut erhoben. Das Ergebnis zeigt, dass bei der Zielgruppe 21% weniger Autofahrten stattfanden und in der Kontrollgruppe 14%. Die 7-prozentige Differenz ist die Netto-Wirkung. Neben den Wirkungen der Maßnahme existieren also noch weitere Einflussgrößen oder Faktoren, die das Verhalten der Eltern und der Schulkinder beeinflussen.

Wichtig für die Qualität von Evaluationen ist, wie gründlich diese weiteren Faktoren beschrieben werden können. Denn nur wenn diese klar dargestellt sind, können die Ergebnisse zuverlässig auf die Maßnahme zurückgeführt werden. Damit steigt der Wert der aus der Evaluation abgeleiteten Schlussfolgerungen. Die Wahl des Evaluations-Designs ist unabhängig von der Anwendungsebene (EU, Land, Region und Gemeinde) und der Komplexität des angesprochenen Themas.

Um eine aussagekräftige Wirkungsevaluation zu planen, muss vor der Durchführung ein Wirkungsmodell aufgestellt werden (siehe Kapitel 2.2 auf Seite 18). Dazu gehört auch, konkrete Hypothesen, also Wirkungsannahmen, zu formulieren. Diese sollten nach Möglichkeit eindeutig definiert werden und messbar sein. Vor der Umsetzung muss geklärt sein, mit welchen Methoden diese erfasst werden können (vgl. Kapitel 2.5 auf Seite 29 und 4.3 auf Seite 52).

Abbildung 2-1 zeigt die empfohlenen Schritte der Wirkungsevaluation, diese werden nachfolgend im Einzelnen vorgestellt.

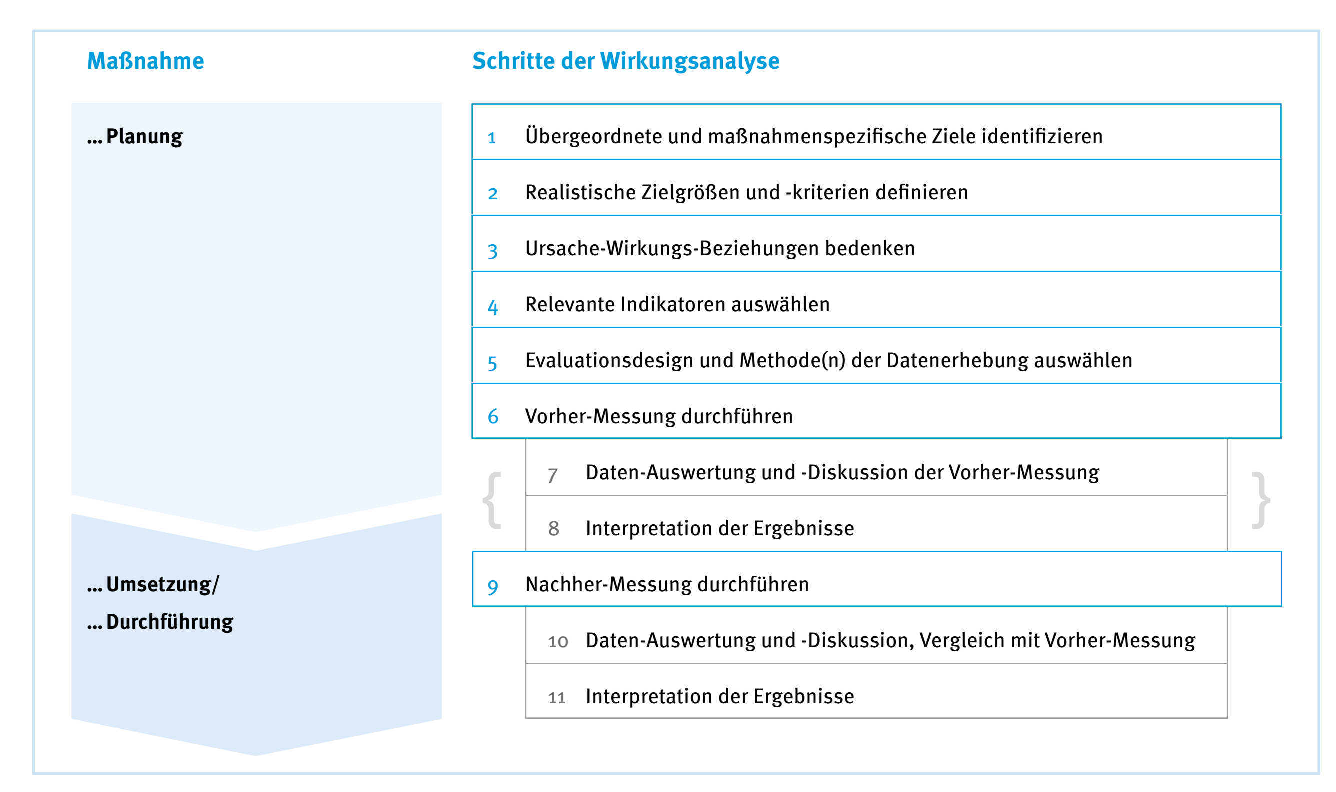


Abb. 2-1: Schritte der Wirkungsevaluation

Oft werden die Daten der Vorher-Messung (Abbildung Schritt 6) zwar zum richtigen Zeitpunkt erhoben, aber deren Auswertung, Ergebnisdiskussion und Interpretation (Schritte 7–8) finden nicht sofort statt, sondern werden auf einen späteren Zeitpunkt verschoben. Das Vorgehen erscheint auf den ersten Blick aus zeitökonomischen Gründen vertretbar. Es ist jedoch empfehlenswert, die oben genannte Reihenfolge strikt einzuhalten, damit es möglich ist noch vor der Maßnahmenumsetzung von den Erkenntnissen zu profitieren und die Maßnahme bei Bedarf entsprechend der gewonnenen Erkenntnisse anzupassen.

2.1 Maßnahmenziele

Im Rahmen einer Evaluation ist oft von Zielen die Rede. Dabei wird der Begriff ganz unterschiedlich verwendet, worauf im Folgenden näher eingegangen wird.

Die kommunale Verkehrsplanung verfolgt einerseits übergeordnete Ziele, so wie sie beispielsweise in einem städtischen Leitbild zu Verkehr und Klimaschutz festgelegt sind. Viele der Forderungen können auf Leitlinien zurückgehen, die 1997 durch die OECD-Konferenz für verkehrspolitisches und

–planerisches Handeln aufgestellt wurden (OECD, 1997). Diese Ziele haben oft einen strategischen Charakter und sind meistens dadurch gekennzeichnet, dass sie einen Idealzustand beschreiben, der oft nicht klar fassbar ist, z. B. »Lebenswerte Städte«, »Förderung des Langsamverkehrs«, »Stadt der kurzen Wege«, »Verringerung der verkehrsbedingten Emissionen«, etc.

Im nächsten Schritt gilt es also, zu den übergeordneten Zielen konkrete und maßnahmenspezifische Teilziele zu definieren. Diese sollten mit Hilfe von einzelnen Maßnahmen erreichbar sein. Ob das Ziel erreicht wurde, sollte sich jeweils mit Hilfe von Indikatoren möglichst zuverlässig und eindeutig messen oder beschreiben lassen. Die Auswahl der realistischen und messbaren Zielgrößen ist nicht immer einfach und sollte daher sorgsam abgewogen werden.

In Evalstadt gelten die übergeordneten Ziele »Förderung des Langsamverkehrs« und »Stadt der kurzen Wege«. Daher werden u.a. als Maßnahme an einer vierspurigen Hauptstraße in der Ortsmitte zusätzliche Querungshilfen für Fußverkehr und Radverkehr in Form einer Bedarfsampel mit Zebrastreifen angelegt. Als Messgrößen gelten a) die Anzahl der Personen, die die Straße in diesem Abschnitt queren sowie b) die Wartezeit vor dem Queren. Der Vergleich der Daten vor und nach dem Anlegen der Querungshilfe zeigt den Effekt der Maßnahme.

An diesem Beispiel kann auch verdeutlicht werden, dass eine gemessene Wirkung nicht nur auf die Maßnahme zurückzuführen ist, sondern dass es oft weitere Einflussgrößen gibt. Jahreszeitliche und wetterbedingte Effekte haben, wie der Wochentag, die Uhrzeit, Feiertage oder Ferien oft einen großen Einfluss auf das Verkehrsverhalten. Vorher- und Nachher-Messung müssen daher unter möglichst ähnlichen Bedingungen stattfinden. Oft ist es im Rahmen von kommunaler Verkehrsplanung üblich, gleichzeitig mehrere Maßnahmen umzusetzen. Wenn also auf der o.g. vierspurigen Straße neben dem neuen Überweg gleichzeitig auch eine neue Geschwindigkeit festgelegt wurde, bspw. statt 50 nur noch 30 Stundenkilometer, dann wird sich diese Änderung ebenfalls möglicherweise auf die Wartezeit auswirken.

Um die tatsächlichen Effekte einer bestimmten Maßnahme möglichst eindeutig zu beschreiben, sollten Sie in jedem Fall eine Kontrollmessung durchführen. Im o.g. Beispiel werden dabei die gleichen Daten auf einem Streckenabschnitt erhoben, an dem keine Querungshilfen existieren.

Das folgende Beispiel der Schulumfeldgestaltung in der Musterstadt Evalstadt beruht in weiten Teilen auf einem Projekt der Stadt Utrecht (Niederlande). Utrecht unterstützt Grundschulen, die die Fragen der Sicherheit im Straßenverkehr proaktiv angehen. Diese Initiative hat einen Einfluss auf das Schulumfeld, den Lehrplan der Schule, und sie fördert die Einbeziehung der Eltern bei sicherheitsrelevanten Fragen.

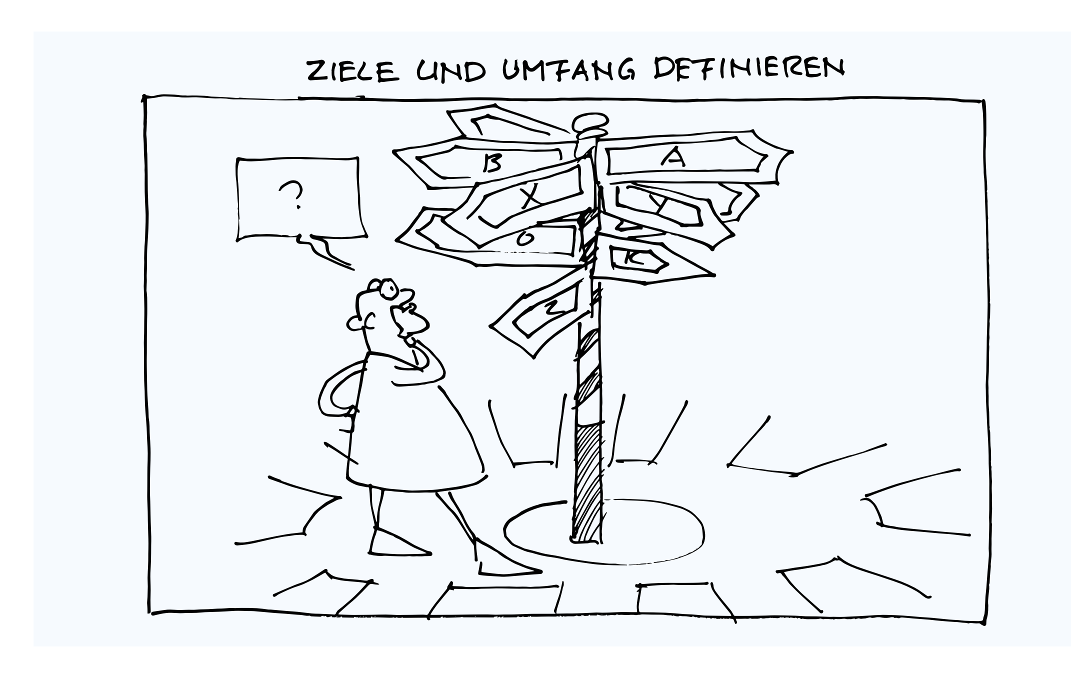
In Evalstadt wurde die verkehrliche Umgebung aller Schulen gleich gestaltet, um die Sicherheit von Fußgängern und insbesondere der Schülerinnen und Schüler zu erhöhen. Zur Umgestaltung gehören Warnschilder, Park- und Halteverbotszonen, Tempo-30-Zonen, auffällig farbige Straßenmarkierungen und die einheitliche Gestaltung der Gehwege. Auf diese Weise soll die Aufmerksamkeit der Autofahrer und -fahrerinnen in der Nähe von Schulen stadtweit erhöht werden, so dass diese ihr Fahrverhalten entsprechend anpassen.

Die vereinheitlichte Schulumgebung wirkt sich auf die durchschnittliche Fahrzeuggeschwindigkeit in der Schulumgebung aus, da die Autofahrenden Geschwindigkeitsbeschränkungen bewusster wahrnehmen und einhalten. Dadurch wird die Aufmerksamkeit insgesamt erhöht, dies wiederum zeigt Wirkungen auf die wahrgenommene und tatsächliche Sicherheit und den Verkehrsfluss. Die Schülerinnen und Schüler werden durch die Schulleitung ermutigt, mit dem Fahrrad oder zu Fuß zur Schule zu kommen, um den Hol- und Bringverkehr durch die Eltern zu minimieren.

Langfristig zielen die eingeführten Maßnahmen darauf ab, die Verkehrsmittelwahl in Richtung nachhaltige Verkehrsträger zu verschieben und damit einen Beitrag zur verbesserten Lebensqualität zu leisten. Diese langfristigen Ziele entsprechen dem Gesamtplan zur urbanen Mobilität, der darauf

abzielt, nachhaltige Verkehrsträger zu stärken und die Sicherheit im städtischen Straßenverkehr zu verbessern.

Für das Formulieren von eindeutigen, maßnahmenspezifischen Zielen ist der SMART-Ansatz hilfreich (Abb. 2-2).



SMART-DIMENSIONEN	Beispiel Einheitliche Schulumgebungen
Spezifisch: Ist das Ziel spezifisch, konkret, exakt, eindeutig, verständlich?	Ein spezifisch festgelegtes Ziel wird viel eher erreicht als ein generelles. Ein spezifisches Ziel ist z. B. die Erhöhung des Radverkehrsanteils der Schülerinnen und Schüler auf den Strecken zwischen Wohnung und Schule
Messbar: Ist es bei diesem Ziel möglich, den Erfolg oder Misserfolg der Maßnahme zu messen? Was ist der Beweis für einen Erfolg?	Um messbar zu sein, muss bei dem konkret formulierten Ziel angegeben sein, was eine deutliche Steigerung bedeutet. Hier beispielsweise 5%. Die Veränderung bei der Verkehrsmittelwahl kann durch eine Erhebung der aktuell genutzten Verkehrsträger vor und nach der Umsetzung gemessen werden. Damit liefert eine Änderung des Modal Split der Stichproben bezüglich Autofahrten, Radfahren und zu Fuß zur Schule gehen quantifizierte Ergebnisse, die dieses Kriterium erfüllen.
Ausführbar: Sind die gesetzten Ziele erreichbar?	Grundvoraussetzung für die Messung der Wirksamkeit einer Maßnahme ist die Bereitstellung bzw. Erhebung von Daten vor der Einführung der Maßnahme. So lässt sich eine Reduktion der Schulfahrten mit dem Auto zugunsten des Fahrrads und des zu Fuß gehen beobachten bzw. messen. Nach einem Blick auf die vorhandenen Daten zum Anteil der Autonutzung für diese Fahrten vor der Maßnahme, stellt man fest, dass das Ziel einer Verschiebung der Hälfte der Stichprobe zu Fahrradfahren und zu Fuß gehen unerreichbar und zu ehrgeizig formuliert ist.

Realistisch: Ist es möglich, das Ziel mit den verfügbaren Ressourcen zu erreichen? Und passt es zu den allgemeinen Zielen?	Es ist nicht hilfreich, zu ehrgeizige Ziele zu formulieren. Basierend auf den Erfahrungen von Fachleuten und den verfügbaren Informationen über ähnliche Projekte in anderen Städten können angemessene Zielgrößen festgelegt werden. In einem geringen Umfang können Ziele auch im Laufe der Maßnahme an starke Veränderungen der Projektbedingungen angepasst werden. Dies sollte jedoch immer auf transparente Art und Weise begründet und kommuniziert werden.
Termingerecht: In welchem Zeitrahmen soll das Ziel erreicht werden?	Innerhalb der Laufzeit des Programms zielt die Stadt auf eine Verringerung des Anteils an Autofahrten auf der Strecke Zuhause-Schule von zirka 5% zugunsten von Rad fahren oder zu Fuß gehen ab. Damit ist der Zeitrahmen klar. Das Ergebnis kann je nach Zeitpunkt der Messung nach der Umsetzung variieren.

Abbildung 2-2: SMART-Dimensionen von Zielen

2.2 Ursache und Wirkung

Fällt das Gewicht auf den Fuß, wenn das Seil losgelassen wird? Der erste Blick auf die Karikatur legt dies nahe. Aber ist das Seil mit dem Gewicht verbunden oder sind es zwei unterschiedliche Seile?

In diesem Abschnitt werden zwei Methoden vorgestellt, um den vermuteten Zusammenhang zwischen Ursache und Wirkung abzuschätzen. Die erste ist die Ursache-Wirkungskette und die zweite eine Variante der Zielgruppenanalyse.



Um alle relevanten Beziehungen zwischen den konkreten Zielen und den möglichen Wirkungen (Leistung, Ergebnis, Auswirkung, siehe Kapitel 2.1 auf Seite 15) zu erkennen, ist das Aufstellen einer *Ursache-Wirkungskette* hilfreich. Sie verdeutlicht die Zusammenhänge zwischen den möglichen Auswirkungen, sowohl positiver als auch negativer Art, und den Ressourcen, die für die Umsetzung der Maßnahme benötigt werden. Das Aufstellen hilft dabei, das breite Spektrum der möglichen, beabsichtigten und unbeabsichtigten, Auswirkungen der Maßnahme zu verstehen und die Wahl der richtigen Indikatoren zu treffen.

Das Ergebnis einer Maßnahme kann stets durch eine Vielzahl von Wirkungen entstehen, die berücksichtigt werden müssen. Oft sind die Auswirkungen auch indirekt. Eine festgestellte Abnahme der Verkehrsunfälle muss nicht unbedingt durch die eingeführte Verkehrsmaßnahme verursacht sein. Auswirkungen und Ergebnisse, sowohl verkehrsbezogen als auch darüber hinaus, und zugeordnete Kosten und Nutzen sind die Elemente der folgenden Illustration zu Ursache und Wirkung. In Abbildung 2-3 wird eine Ursache-Wirkungskette für das Beispielprojekt »Einheitliche Schulumgebung« vorgestellt, um die denkbaren Auswirkungen auf unterschiedliche Handlungsfelder exemplarisch anzugeben. Dargestellt wird auch die Richtung der Wirkungsbeziehung.

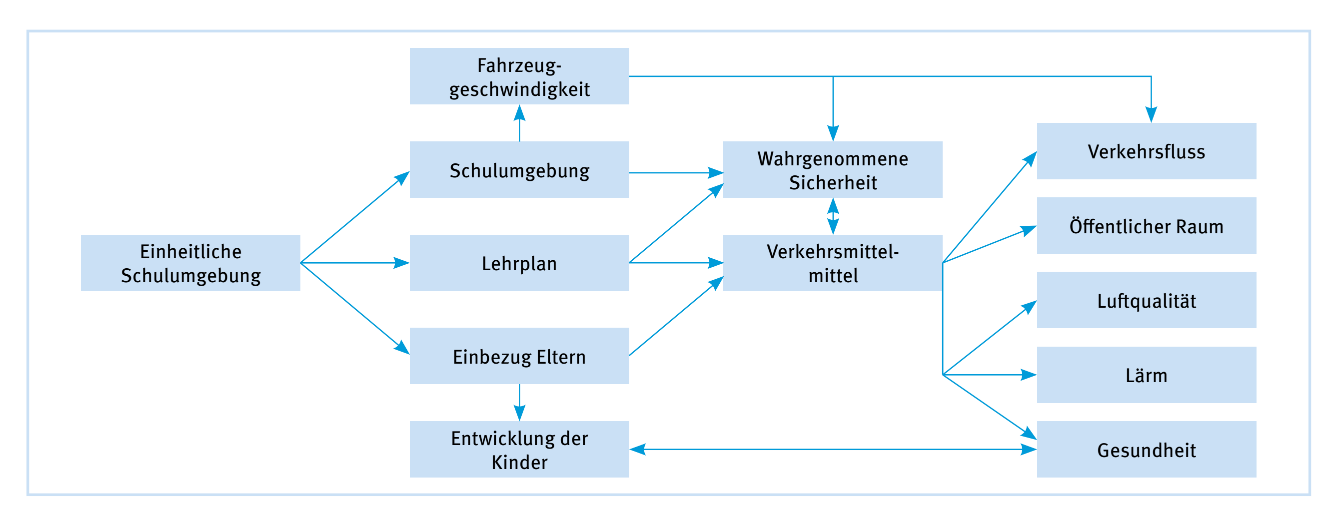


Abb. 2-3: Ursache-Wirkungskette für das Beispiel Einheitliche Schulumgebung

Im Folgenden werden mögliche Ursachen und deren Wirkungen vorgestellt. Die Einbeziehung von Mobilitätsbildung in den Lehrplan der Schulen könnte Auswirkungen auf die Verkehrsmittelwahl haben, da möglicherweise mehr Eltern ihren Kindern erlauben, mit dem Fahrrad oder zu Fuß zur Schule zu kommen. Dies führt zu mehr Sicherheit und einer entsprechenden Wahrnehmung sowie zu Veränderungen beim Verkehrsfluss, der Nutzung des öffentlichen Raums, der Luftqualität und der Gesundheit der Kinder. Die bessere Gesundheit könnte auch durch eine veränderte Einbeziehung der Eltern gefördert worden sein. Im Programm werden Eltern auch dazu ermutigt, ihren Kindern beizubringen, selbständig verantwortungsvoll zu handeln, was deren persönliche Entwicklung unterstützt.

Neben den erwarteten positiven Auswirkungen sollten im Rahmen eines Wirkungsmodells auch mögliche negative Folgen bedacht werden. Eine Neuordnung von Verkehrszeichen und Fahrbahnmarkierungen in der Schulumgebung bringt eventuell nicht nur die gewünschte Entwicklung, sondern könnte auch dazu führen, dass Schulkinder im Straßenverkehr insgesamt weniger aufmerksam sind.

Lärmpegel, Luftqualität, die Nutzung des öffentlichen Raums sowie die Gesundheit der Kinder sind das Ergebnis zahlreicher Faktoren. Beispielsweise könnten Schulen mit einem sportlichen Schwerpunkt bereits »gesündere« Schüler haben. Aus diesem Grund wurden diese Faktoren und deren Entwicklung von der Analyse ausgeschlossen. Aus dem gleichen Grund wird auch der Verkehrsfluss nicht evaluiert. Die vereinfachte Ursache-Wirkungskette ist in der nächsten Abbildung (Abbildung 2-4) zu sehen.

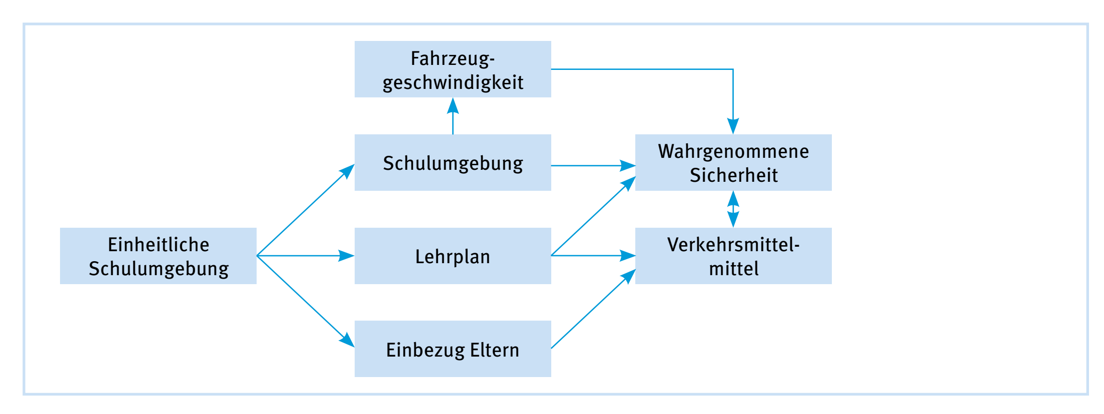


Abb. 2-4: Kondensierte Ursache-Wirkungskette für das Beispiel Einheitliche Schulumgebung

Damit konzentriert sich die eigentliche Evaluation auf die Geschwindigkeit der Fahrzeuge in der Schulumgebung, die wahrgenommene Sicherheit der Kinder und die Verkehrsmittelwahl. Für jedes Ziel sollte mindestens eine Auswirkung formuliert werden. Die Indikatoren müssen passend zu den Zielen ausgewählt werden (siehe Kapitel 2.3 auf Seite 21). Wenn die Ursache-Wirkungskette darauf hinweist, dass die Ziele schwer oder gar nicht zu erreichen sind, können die Ziele an dieser Stelle noch einmal korrigiert werden.

Eine zweite Methode für die Abschätzung der Wirkung ist die *Zielgruppenanalyse*. Um die Beziehung zwischen Ursache und Wirkung abzuschätzen, wird eine Liste aller relevanten Personen oder Gruppen aufgestellt. Für jede der Gruppen werden Vor- und Nachteile, Kosten und Nutzen sowie erwartete negative und positive Auswirkungen beschrieben (siehe Abbildung 2-5).

Schülerinnen und Schüler	+ weniger Unfälle, verringertes Unfallrisiko, höhere Sicherheit + werden ermutigt, die Schulwege zu Fuß oder mit dem Rad zurückzulegen + höhere Verantwortung, geringere Abhängigkeit von den Eltern – könnten annehmen, dass sich Autofahrer auch außerhalb der Schulumgebung rücksichtsvoll verhalten
Lehrerinnen und Lehrer	+ Wahrnehmung einer sicheren Schulumgebung + höhere Verantwortung für Mobilitätserziehung
Eltern	+ Überdenken ihre Autonutzung – haben weniger Kontrolle darüber, ob Kinder wohlbehalten zur Schule und zurück kommen, Unsicherheit/Sorge
Verwaltung	+ Imagesteigerung bei erfolgreicher Umsetzung – Finanzierung der Maßnahmen
Autofahrerinnen und Autofahrer	+ erhöhte Aufmerksamkeit gegenüber Schülerinnen und Schülern und Zufußgehenden + gesteigerte Wahrnehmung von Schulumgebungen, geringeres Unfallrisiko – Verunsicherung wegen ungewohnten und neuen Markierungen und Zeichen
Nachbarn	+ geringeres Unfallrisiko für sich und besonders für Kinder + regulierter, langsamer Verkehr + bessere Luftqualität, bessere Lebensqualität

Abbildung 2-5: Wirkungen auf Interessengruppen – Beispiel Schulumgebung

Das Aufstellen einer solchen systematischen Übersicht über alle von der Maßnahme betroffenen Gruppen sowie das Zuordnen von möglichen positiven und negativen Effekten hilft dabei, die erwünschten Wirkungen und die unerwünschten »Nebenwirkungen« auf der Ebene der Menschen in ihrer Rolle als Verursacher, Betroffene oder Handelnde besser abzuschätzen. Die Methode ergänzt daher die rein technische oder bauliche Perspektive sehr gut. Mit Hilfe dieser Liste können auch weitere Maßnahmen geplant werden, wie zum Beispiel eine begleitende Informationskampagne zur Einführung der Maßnahme, Handlungs- oder Gebrauchsanweisungen, ein Anreizsystem zum Mitmachen, die Kommunikation zum Rückmelden der Erfolge und Effekte der Maßnahme usw.

Das Aufstellen einer Ursache-Wirkungskette und die Analyse der Zielgruppen ist keine lästige Übung oder gar Zeitverschwendung, sondern wichtiger Bestandteil einer Evaluation. Sie kommen auch nach der Analyse der Ergebnisse wieder zum Einsatz, da sie bei der Interpretation der Ergebnisse und zur Planung weiterer Maßnahmen hilfreich sind.

2.3 Indikatoren

Nach dem Festlegen der Ziele und dem Beschreiben der Ursache-Wirkungsbeziehungen werden im dritten Schritt die wichtigsten Indikatoren ausgewählt. Indikatoren sind Messgrößen, die mögliche Wirkungen der Maßnahme unter dem gegebenen Budget am besten nachweisen.

Bei der Auswahl von Indikatoren sollten drei grundlegende Anforderungen berücksichtigt werden:

- Die Leistung oder Wirkung der Maßnahme muss damit erfasst werden können (siehe Kapitel 2.2 auf Seite 18),
- sie müssen zu den Zielen passen (siehe Kapitel 2.1 auf Seite 15) und
- es muss Methoden oder Verfahren geben, mit denen sie verlässlich gemessen werden können (siehe Kapitel 2.5 auf Seite 29).

Es gibt verschiedene Einzelindikatoren oder auch Indikatorensets, die im Rahmen von deutschen oder europäischen Verkehrsprojekten entwickelt worden sind. Diese können hilfreich für die Auswahl maßnahmenspezifischer Indikatoren sein. Oft liegen sogar Daten für diese Messgrößen vor, die Sie im Rahmen Ihres Projektes zur Einordnung nutzen können. Aber nicht jedes Indikatorenset passt auf jede Maßnahme, ggf. muss man dieses anpassen. Der Einsatz hängt auch davon ab, ob die benötigten Messgrößen in Ihrer Kommune erhoben werden können oder sogar schon vorliegen. Besonders sinnvoll ist der Einsatz von Indikatoren, welche auch in anderen Städten mehr oder weniger regelmäßig erhoben werden, denn so können Daten und Programme gegenübergestellt werden (siehe auch Kapitel 10 auf Seite 80).

Innerhalb eines Programms können Indikatoren nach Kontext und ihrer Wichtigkeit festgelegt werden, beispielsweise in indikatorenspezifischen Methodenblättern.

Versuchen Sie beispielsweise, Alternativen zu individuell genutzten motorisierten Fahrzeugen in Ihrer Stadt durch Verkehrsverlagerung zu fördern und zu messen, sollten Sie die Verkehrsmittelwahl anschauen. Die Daten können durch den quantitativen Messwert durchschnittliche Verkehrsmittelwahl (Passagiere, Fahrzeuge oder Fahrten) erhoben werden. Dieser kann die Anteile der Passagiere, Fahrzeugkilometer oder der Fahrten für jedes Verkehrsmittel im Jahresverlauf beschreiben. Der Anteil nicht-motorisierter Verkehrsmittel, wie beispielsweise zu Fuß gehen und Radfahren, kann mit dem Anteil von öffentlichen Verkehrsmitteln (Bus, Straßenbahn, U-Bahn, Zug) und dem motorisierten Individualverkehr mit dem Auto oder Motorrad verglichen werden.

Bitte beachten Sie, dass für den bestmöglichen Einblick in die Auswirkungen Ihrer Maßnahmen möglicherweise weitere Interpretationen erforderlich sind und die Maßnahmen möglicherweise noch um zusätzliche lokale Indikatoren ergänzt werden müssen (siehe Kasten 2).

Kasten 2: Allgemeine und lokale Indikatoren

Besonders im Rahmen größerer EU-Projekte werden oft Indikatoren-Listen erarbeitet, beispielsweise die Maestro-Indikatorliste. Es lohnt sich, vorhandene Indikatorenlisten daraufhin zu prüfen, ob diese auch auf Ihre Maßnahme passen, denn sie werden bei einer standortübergreifenden Evaluation helfen, also den Vergleich mit anderen ermöglichen. Es ist aber eher unwahrscheinlich, dass Sie Listen finden, die vollständig zur Messung der Wirkungen in Ihrem Projekt ausreichen. Sie werden daher wahrscheinlich weitere, spezifische lokale Indikatoren für Ihre Evaluation finden müssen. Diese Indikatoren können verwendet werden:

- um eine Bewertung auf einer detaillierten Ebene durchzuführen,
- um die Auswirkungen in Bezug auf ein bestimmtes lokales Problem zu beurteilen, und
- um die Auswirkungen der Besonderheiten einer lokalen Maßnahme zu beurteilen.

Im Beispiel »Schulumgebung« führen die Ziele und die Ursache-Wirkungskette zu folgenden Indikatoren:

Spezifisches Ziel	Indikator	Beschreibung der Indikatoren
»Die Zufriedenheit mit der Verkehrssicherheit bei Kindern, deren Eltern und Lehrpersonal um 30% erhöhen«	Wahrnehmung der Sicherheit	Veränderung bei der Wahrnehmung der Verkehrssicherheit bei Lehrkräften, Schulkindern und Eltern
»Die Zahl der Unfälle mit Kindern in der Umgebung von Schulen um 5% verringern«	Sicherheit	Anzahl der Kinderunfälle in der Umgebung von Schulen
»Den Anteil der Autofahrten zur Schule um 5% zugunsten der Verkehrsmittelwahl Fahrrad und zu Fuß gehen verringern.«	Verkehrsmittelwahl	Durchschnittliche Prozentzahl der Strecke Schule-Zuhause für jedes Verkehrsmittel

Tabelle 2-1: Liste der möglichen Indikatoren für das Beispiel Schulumgebung

Zur Erhebung der wahrgenommenen Sicherheit der Schulumgebung wird eine schriftliche oder mündliche Befragung der Eltern, Lehrkräfte und Schulkinder durchgeführt. Zur Bewertung der Maßnahme müssen die Anzahl der Fahrten nach Verkehrsmittel erhoben werden, bevor die Maßnahme umgesetzt wird (Referenz) und danach (für die Wirkungsevaluation). Die Differenz zwischen den Anteilen der Verkehrsmittel vor und nach der Umsetzung, abzüglich der Änderungen, die auch ohne Maßnahme eingetreten wären, beschreiben die Wirkung der umgesetzten Maßnahme (siehe Tabelle 2-1.).

Wenn gleichzeitig mehrere Verkehrsmaßnahmen parallel evaluiert werden, sollte kritisch geprüft werden, ob es Indikatoren gibt, die von mehr als einer Maßnahme beeinflusst werden könnten. Solche Maßnahmen sind »gebündelte Maßnahmen« mit sogenannten »gebündelten Indikatoren« (siehe Kasten 3).

Kasten 3: Gebündelte Indikatoren

Wenn Sie mehrere Maßnahmen in Ihrem Projekt haben und Sie ihre Ursache-Wirkungsketten vergleichen, möchten Sie vielleicht eine oder mehrere Indikatoren wählen, die für mehr als eine Maßnahme sinnvoll und anwendbar sind. Die Wirkung, die durch diese Messgröße festgestellt werden, könnte auch durch andere Maßnahmen beeinflusst worden sein, als Ergebnis der kombinierten Wirkung von zwei oder mehr Maßnahmen. Dabei handelt es sich um maßnahmenübergreifende und nicht um maßnahmenspezifische Indikatoren. Idealerweise sollten die Indikatoren für eine Maßnahme nicht nur aus gebündelten Indikatoren bestehen. Dies würde es unmöglich machen, eine klare Aussage über den Erfolg oder Misserfolg dieser speziellen Maßnahme zu treffen, da die Wirkung immer auch einer anderen Maßnahme zugeschrieben werden könnte.

Viele andere Eingriffe – innerhalb Ihrer Maßnahmenkombination oder durch externe Einflüsse – können eine Messgröße beeinflussen (z. B. die Maßnahme an sich, siehe Referenz-Szenario, Kapitel 2.4 auf Seite 23). Natürlich können nicht alle Einflüsse berücksichtigt und gemessen werden, aber die wichtigsten sollten in der Interpretation der Ergebnisse angesprochen werden.

2.4 Untersuchungsdesigns

Im Untersuchungsdesign oder Erhebungsdesign wird festgelegt, wo, wann und wie oft und mit welchen technischen, personellen und finanziellen Rahmenbedingungen die Evaluation durchgeführt wird, um die nötigen Daten zu sammeln. Die Auswahl eines bestimmten Designs wird häufig durch notwendige Abwägungsprozesse zwischen Kosten und der Zuverlässigkeit der Ergebnisse beeinflusst. Allgemein gilt: Je mehr Sicherheit Sie bezüglich der Ergebnisse und Auswirkungen Ihrer Maßnahmen benötigen, desto aufwändiger wird die Evaluation.

In vielen Fällen stehen bereits ausreichende Daten aus verschiedenen Quellen zur Verfügung und können in das Evaluationsdesign einbezogen werden (siehe Kapitel 10 auf Seite 80). Diese müssen analysiert und mit den Maßnahmenzielen und der Wahl der Indikatoren (z. B. Unfalldaten der Polizei, Daten zur Verkehrsdichte von den Straßenplanungsämtern, Daten zur Luftqualität von den Umweltbehörden) verknüpft werden. Es gibt deutliche Unterschiede der Evaluationsdesigns zwischen:

- Designs mit randomisierten Kontrollgruppen,
- experimentellen Designs mit Vergleichsgruppe und
- experimentellen Designs ohne Vergleichsgruppe.

Diese werden im nächsten Abschnitt genauer beschrieben. Neben der Planung des Evaluationsdesigns ist es notwendig, die Indikatoren festzulegen, für die vorher und nachher gemessen werden soll. Wichtig ist die Erhebung der notwendigen Daten für alle relevanten Indikatoren vor der Umsetzung der Maßnahme. Diese werden dann mit den Daten der Erhebung nach der Umsetzung der Maßnahme verglichen. Um aussagekräftige Auswertungsergebnisse zu erhalten, müssen Daten für das gesamte Indikatorenset vorliegen. Außerdem sollte ein ausreichend großer Datensatz vorliegen, um die zu erwartenden Veränderungen statistisch auswerten zu können, wenn dies sinnvoll und möglich ist.

Im Beispiel Schulumgebung in Evalstadt müssen daher für jeden der Standorte, an denen die einheitliche Schulumgebungen umgesetzt werden, bereits im Vorfeld Daten zur Sicherheitswahrnehmung der Eltern, den tatsächlichen Unfällen und die Verkehrsmittelwahl der Schulkinder erhoben werden. Wichtig hierbei ist, dass diese Vorherdaten auch für den Kontrollstandort erhoben werden.

Idealerweise werden an Kontrollstandorten, also da wo die Situation ganz vergleichbar ist, aber keine Maßnahmen durchgeführt werden, die gleichen Arten von Daten zum gleichen Zeitpunkt wie am Maßnahmenstandort erfasst. Aus dem Kontrollstandort ergibt sich das so genannte Referenzszenario, also die Entwicklung der Indikatoren, ohne Maßnahme (Abb. 2-6). Ein solches Szenario kann aber

neben dem Vergleich einer Kontroll- oder Vergleichsgruppe oder eines Kontroll- bzw. Vergleichstandorts auch mit Hilfe von historischen Daten oder einer Modellierung aufgestellt werden.

Kasten 4: Wozu dienen Kontroll- oder Vergleichsgruppen?

Neben den im Projekt durchgeführten Maßnahmen können viele weitere Faktoren das Erleben und Verhalten der Menschen beeinflussen. Um diese Faktoren und ihre Wirkung angemessen zu berücksichtigen, erhebt eine gute Evaluation auch immer die Daten einer Kontrollgruppe an einem Kontrollstandort oder entsprechend einer Vergleichsgruppe.

Im Fall Evalstadt wird also eine weitere Schule mit möglichst ähnlichen Bedingungen ausgewählt, an der keine verkehrsbezogenen Änderungen der Infrastruktur und keine Kommunikationsmaßnahmen durchgeführt werden. An diesem Kontrollstandort werden die gleichen Fragebögen zur gleichen Zeit wie am Standort der Maßnahme verteilt werden, also auch vorher und nachher. Der Unterschied zwischen den Werten des Standorts mit durchgeführter Maßnahme und dem Kontrollstandort ist die eigentliche Wirkung der Maßnahme, da andere Einflüsse, beispielsweise Benzinpreisschwankungen oder Medienaufmerksamkeit für das Thema Schulwegeunfälle auch am Kontrollstandort aufgetreten wären.

Oft ist auch schon die Datenerhebung an sich, beispielsweise die Befragung, eine Art Maßnahme oder Eingriff und könnte einen Einfluss auf das Verhalten zeigen. Dies ist bei der Wahl des Termins und der Kommunikation zum Grund der Befragung zu berücksichtigen.

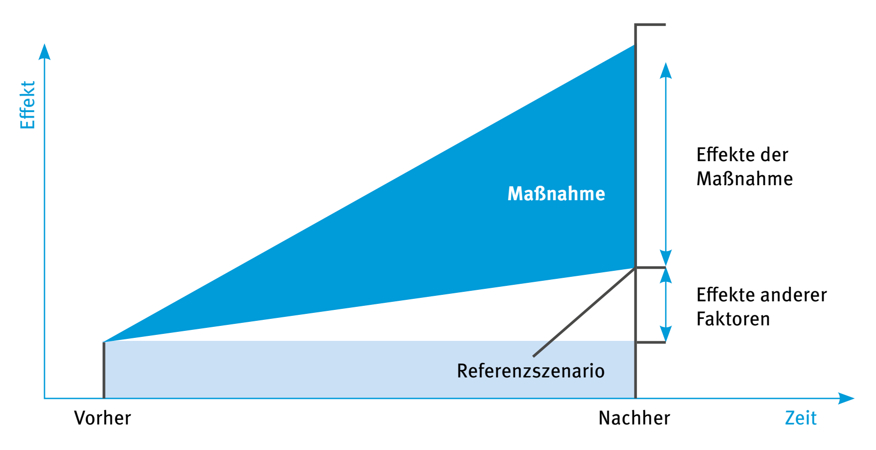


Abbildung 2-6: Die Bedeutung des Referenzszenarios

Randomisierte Kontrollgruppen

Im Idealfall erfolgt die Zuordnung zu Experimental-, Vergleichs- oder zur Kontrollgruppe nach dem Zufallsprinzip. Dazu kann eine Tabelle von Zufallszahlen oder eine rechnergestützte Zufallsauswahl verwendet werden. Wichtig ist, dass ein Vergleich möglichst gleichartiger Gruppen (z. B. in Bezug auf Alter, Geschlecht und Beruf) gewährleistet ist. Dann sind die Unterschiede, die zwischen diesen Gruppen beobachtet werden, wahrscheinlich(er) auf die Maßnahme zurückzuführen, und nicht auf Unterschiede zwischen den Gruppen, die bereits zu Beginn der Studie vorhanden waren. Gleiches gilt auch für die Zuordnung von Maßnahmenstandorten (z. B. zwei möglichst ähnliche Straßen oder Schulen oder Kreuzungen). Das randomisierte Experiment wird oft als Goldstandard bezeichnet, weil es den höchsten Qualitätskriterien gerecht wird.

Beide Gruppen oder Standorte werden vor und nach der Umsetzung der Maßnahme mit den gleichen Methoden zur Datenerhebung (z. B. Umfragen, Messungen) evaluiert. Die Ergebnisse der Kontrollgruppe oder des Kontrollstandorts entsprechen dem Referenzszenario.

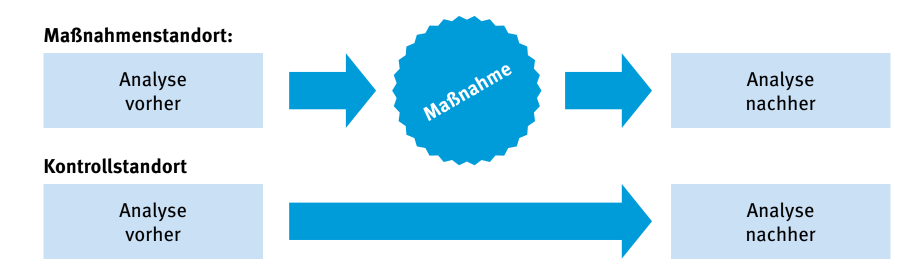


Abbildung 2-7: Datenerfassung Vorher-Nachher beim Maßnahmenstandort und einem Kontrollstandort

Quasi-experimentelle Vergleichsgruppen

Bei vielen Gegebenheiten im Alltag einer Kommune ist eine zufällige Zuordnung nicht möglich. Oft geht es ja gerade darum, eine nur an einem Standort geplante und umgesetzte Maßnahme zu evaluieren. Da ist es nicht möglich, auszuwürfeln, an welchem Standort sie umgesetzt werden soll. In einem solchen Fall muss man sich darum bemühen, einen Standort oder eine Personengruppe zu finden, die der ersten in möglichst vielen Punkten möglichst ähnlich ist. Dieses Untersuchungsdesign wird als Quasi-Experiment bezeichnet.

Der Begriff Quasi-Experiment bezieht sich auf ein Forschungsdesign, das viele Ähnlichkeiten mit der randomisierten, kontrollierten Studie hat, aber es fehlt speziell das Element der zufälligen Zuweisung. Bei einer Zufallszuweisung haben alle Teilnehmenden die gleiche Chance, an einer Maßnahme teilzunehmen. Dadurch stellt die Zufallszuweisung sicher, dass sowohl Experiment- als auch Kontrollgruppe gleichwertig sind.

Im Projekt Schulumgebung in Evalstadt würde eine Zufallszuweisung bedeuten, dass die teilnehmenden Schulkinder, Eltern und Lehrkräfte nach dem Zufallsprinzip einer Schule zugewiesen werden. Dies ist selbstverständlich nicht möglich, das Projekt arbeitet mit den bestehenden Schulklassen.

Die Schulen nach dem Zufallsprinzip auszuwählen, wäre ebenfalls quasi-experimentell, da die Faktoren, beispielsweise das Gebiet, in dem sich die Schule befindet, nicht »kontrolliert« werden. Daher sind die meisten Forschungsdesign für Verkehrsmaßnahmen tatsächlich quasi-experimentell.

Die quasi-experimentellen Vergleichsgruppen und -standorte können sich von den Maßnahmengruppen und -standorten in vielen Aspekten systematisch unterscheiden. Dies gilt es bei der Auswahl sorgfältig zu beachten, denn es könnten mögliche Erklärungen für Unterschiede bei den Effekten sein. Daher ist es wichtig, alle relevanten Variablen zu beschreiben und ggf. zu untersuchen.

Im Evalstädter Projekt »Schulumgebung« nehmen zunächst zwei Schulen teil, die in einem Wohngebiet an einer mäßig stark befahrenen Straße mit breiten Geh- und Radwegen liegen. Die Lehrerschaft und Eltern sind engagiert und aufgeschlossen für die Themen Langsamverkehr und Sicherheit. Die beiden Schulen, die als jeweilige Vergleichsstandorte ausgewählt und zur Teilnahme gewonnen werden können, sind zwar hinsichtlich der Merkmale Architektur, Größe, Schülerzahl, pädagogische Ausrichtung und Altersmischung sehr ähnlich, sie liegen aber eher am Stadtrand mit einer stärker gewerblich ausgerichteten Infrastruktur, etwas mehr Lieferverkehr und es gibt keine Radwege in der unmittelbaren Nähe. Das Einzugsgebiet der Schule ist etwas größer, d.h. die Schüler haben durchschnittlich einen Kilometer weitere Anfahrtswege. Die Lehrer sind engagiert und am Thema interessiert, haben aber aufgrund von einigen Krankheitsfällen deutlich weniger Zeit als die Lehrkräfte am Maßnahmenstandort.

Wenn also die Effekte am Maßnahmenstandort deutlich positiver ausfallen als am Vergleichsstandort, dann kann dies nicht nur auf die Maßnahme selbst zurückgeführt werden. Andere Einflussgrößen, wie z. B. das Einzugsgebiet oder das Fehlen des Radweges können ebenso negative (oder auch positive) Effekte zeigen. In der medizinischen und psychologischen Forschung versucht man,

diese Effekte durch Studien mit eineiigen Zwillingen zu minimieren oder zu kontrollieren. Im Alltag der Verkehrsplanung gibt es keine Zwillinge – aber für die Auswahl von Vergleichsstandorten gilt es, möglichst große Ähnlichkeiten zu finden und die wesentlichen Unterschiede kritisch im Blick zu behalten.



Untersuchungsdesigns ohne Vergleichsgruppen

Kann keine Kontroll- oder Vergleichsgruppe ermittelt oder nur durch hohe Kosten in eine Evaluation eingebunden werden, sind auch »Vorher-Nachher-Vergleiche« und »Panel-Designs« zulässig. In solchen Fällen können aber keine Referenzszenarien, sondern nur die Bruttowirkungen bestimmt werden.

Bei einem Panel-Design wird die gleiche Stichprobe von Teilnehmenden mindestens zweimal befragt oder in ihrem Verhalten beobachtet. Werden die gleichen Personen – das Panel – zweimal befragt, sind sie hinsichtlich aller Merkmale vergleichbar. Allerdings muss das Panel sehr sorgfältig geplant werden, es sollte Personen aller relevanten Zielgruppen in ausreichender Anzahl einschließen. Bei der gewählten Anzahl sollte von vornherein berücksichtigt werden, dass Personen ihre Teilnahme abbrechen können. Dies ist umso eher der Fall, je länger die beiden Untersuchungszeiträume auseinander liegen oder umso aufwändiger die erste Untersuchung war. Um diese so genannte Abbruchrate zu minimieren, könnten entsprechende Anreize oder Belohnungen gesetzt werden. Die Abbruchrate wird jedoch wahrscheinlich nie bei Null liegen. Ein Panel hat eine Reihe von Vorteilen, beispielsweise eine hohe Qualität und Gültigkeit der Daten, es müssen aber auch die hohen Kosten für die Durchführung einkalkuliert werden.

Entscheidung für ein Evaluationsdesign

Bei der Wahl eines Evaluationsdesigns sind das Vorhandensein und die Eigenschaften einer Kontrollgruppe und eines –standortes (oder entsprechender Vergleichsgruppen) sowie die Häufigkeit der Messung wichtige Kriterien. Dies ist unabhängig von der Frage, ob neue Daten erhoben werden oder verfügbares Material analysiert wird.

Die Datenerhebung ist oft ein großer Posten im Budget, sie kann für über die Hälfte der Kosten einer Evaluierung verantwortlich sein und ist zudem eine zeitaufwändige Aufgabe. Daher stellt dieser Abschnitt Möglichkeiten vor, um das Evaluationsdesign zu vereinfachen. Leider gehen die wirtschaftlichsten Designs zu Lasten der Qualität, nachfolgend beispielsweise die Designs 4 und 5. Dennoch können sie nützliche Erkenntnisse für die Praxis liefern.

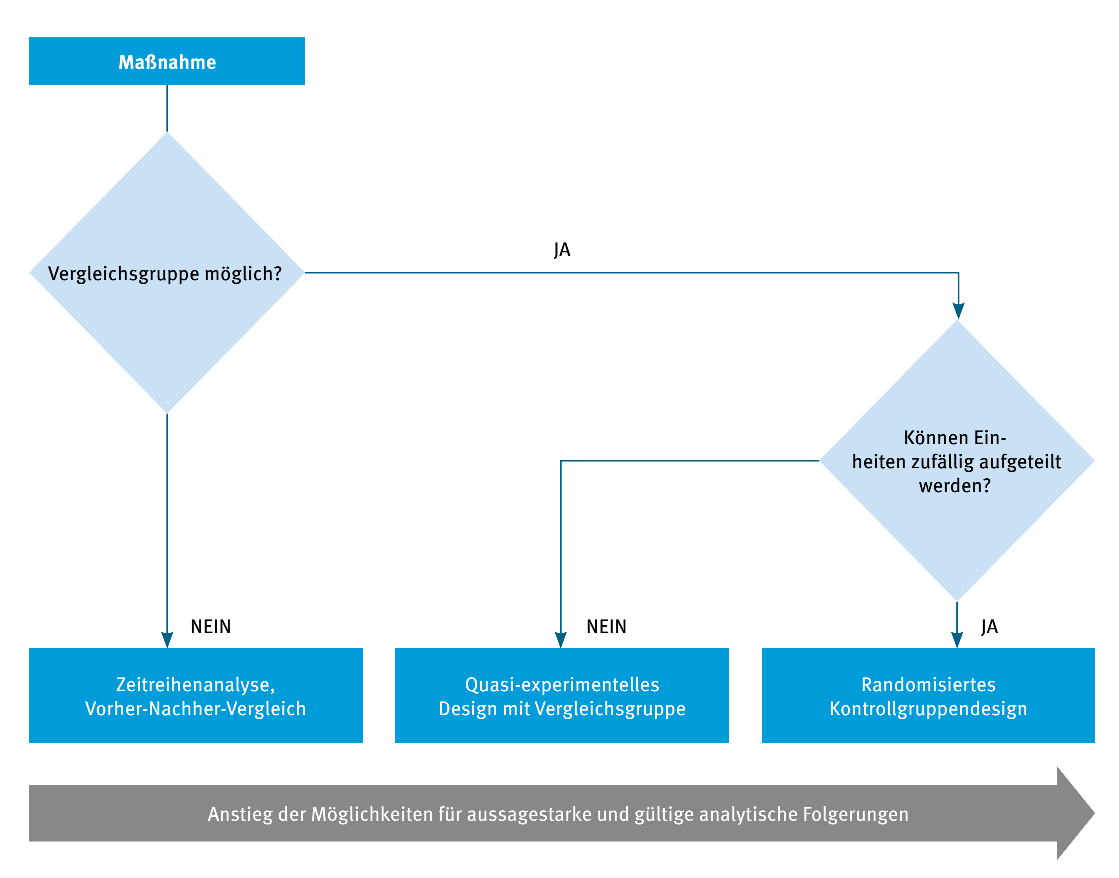


Abbildung 2-8: Auswahl des Evaluationsdesigns

Alle im Folgenden beschriebenen »vereinfachten« Designs verzichten auf eine oder mehrere der folgenden Datenerhebungen: Erhebung der Vorherdaten vor dem Test, Erfassen einer Kontroll- bzw. Vergleichsgruppe vor oder nach der Maßnahme. Jede dieser Verkürzungen mindert die Aussagekraft der Evaluation. Es muss klar gesagt werden, dass die Verwendung einer einfachen Zeitreihenanalyse oder eines Vorher-Nachher-Vergleichs keine präzisen Rückschlüsse auf die Wirkungen zulassen, die ausschließlich durch die Maßnahme entstanden sind. Die vier vereinfachten Designoptionen sind neben dem Goldstandard Design 1 in der nachstehenden Tabelle 2-2 (auf der nächsten Seite) zusammengefasst.

* * * * *	* * * *	* * *	* *	*
Design 1 – das beste	**Design 2**	**Design 3**	**Design 4**	**Design 5**
Datenerhebung in Projekt-und Kontrollgruppe vor und nach der Umsetzung	Datenerhebung in Projektgruppe vor und nach der Umsetzung, Kontrollgruppe nur nach der Umsetzung	Datenerhebung in Projekt- und Kontrollgruppe nur nach Umsetzung; keine Vorherdaten	Datenerhebung nur in Projektgruppe vor und nach Umsetzung; ohne Kontrollgruppe	Datenerhebung in Projektgruppe nach Umsetzung ohne Vorherdaten und ohne Kontrollgruppe
Maßnahmestandort: Analyse vorher → Maßnahme → Analyse nachher **Kontrollstandort:** Analyse vorher → Analyse nachher	**Maßnahmestandort:** Analyse vorher → Maßnahme → Analyse nachher **Kontrollstandort:** Analyse nachher	**Maßnahmestandort:** Maßnahme → Analyse nachher **Kontrollstandort:** Analyse nachher	**Maßnahmestandort:** Analyse vorher → Maßnahme → Analyse nachher	**Maßnahmestandort:** Analyse nachher
Das stärkste Design, da die Evaluation von Beginn an berücksichtigt wurde und zur gleichen Zeit wie die Projektplanung begann (=Goldstandard)	Schwächer als Design 1, aber ermöglicht immerhin einen Vergleich vorher – nachher in der Projektgruppe und einen Nachhervergleich mit den Kontrollgruppendaten	Ein übliches Design, wenn die Evaluation in einem späten Stadium des Projektzyklus oder nach Ende des Projekts beginnt. Durch das Fehlen der Vorherdaten ist es schwierig, die anfänglichen Unterschiede zwischen beiden Gruppen zu kontrollieren	Ein übliches Design, wenn die Daten nur für die Projektgruppe erhoben werden. Methodisch schwach, da die Verwendung der Vorher-Daten als Vergleichsobjekt Annahmen über Wirkungen zu unterschiedlichen Zeitpunkten und individuell nicht beobachtbare Variablen benötigt	Schwächstes Design! Wird jedoch häufig verwendet, wenn die Evaluation spät im Projekt mit sehr begrenzter Zeit und begrenztem Geld beginnt

Tabelle 2-2: Übersicht über mögliche Evaluationsdesigns

2.5 Methoden der Datenerhebung

Als Datengrundlage können grundsätzlich eigene, d. h. primäre, und fremde, d. h. sekundäre Informationen dienen. Primärdaten werden spezifisch für die Evaluation gesammelt, ausgewertet und interpretiert. Werden für die Evaluation bereits vorhandene Daten verwendet oder neu analysiert, ist das eine Sekundärdatenanalyse.

Bei der Planung der Evaluation ist es zeit- und kostensparend zunächst zu schauen, auf welche Sekundärdaten zurückgegriffen werden kann. Das können beispielsweise sein: regelmäßige Verkehrszählungen und Geschwindigkeitsmessungen, Unfallstatistiken, Modal-Split, Anzahl zugelassener Kraftfahrzeuge in einem bestimmten Gebiet, Luftqualität, Passantenströme, Fahrgastzahlen im öffentlichen Nahverkehr, Anzahl der Ticketverkäufe, Fahrgastbefragungen und so weiter.

Datenquellen für eine Evaluation	
Primärdaten (eigens für die Evaluation erhobene Daten)	**Sekundärdaten** (bereits vorliegende Daten)
Werden erhoben durch: · Narrative Interviews · Fragebogenerhebung · Fokusgruppengespräche · Beobachtungen · (...)	**Dazu gehören:** · Unfallstatistiken · regelmäßige Verkehrszählungen · Neu- und Gebrauchtwagenkäufe · Falschparker-Statistik · (...)

Abbildung 2-9: Mögliche Datenquellen für die Evaluation

So verlockend es klingt, auf vorhandenes Material zurückzugreifen, so schwierig kann es auch sein. Es muss sichergestellt werden, dass die Sekundärdaten für die Fragestellung wirklich relevant und zuverlässig sind. Nur weil sie verfügbar sind, sind sie noch lange nicht aussagekräftig. Zudem bedeutet »verfügbar« nicht immer »kostenlos«, in diesem Fall ist die Prüfung, wie sinnvoll sie sind, noch wichtiger.

Im Allgemeinen können Mobilitätsdaten für den Personentransport und den Gütertransport erhoben werden. Für diese beiden Verkehrsarten können die erhobenen Mobilitätsdaten entweder verhaltensbezogen und/oder verkehrsbezogen sein. Verhaltensbezogene Daten beziehen sich immer auf das Erleben und Verhalten von Menschen. Die Informationen können mit Hilfe von Befragungen oder von Beobachtungen erhoben werden. Verkehrsbezogene Daten werden meistens durch Zählungen oder Messungen erhoben (z. B. Fahrzeugzählungen, Geschwindigkeitsmessungen).

Jetzt wird schwerpunktmäßig die Erhebung verhaltensbezogener Daten mit Hilfe unterschiedlicher Befragungstechniken vorgestellt. In der Regel sind das schriftliche oder mündliche Interviews mit ausgewählten Personen auf eine mehr oder weniger strukturierte Art, entweder persönlich oder über Telefon oder Internet. Im Folgenden werden die verschiedenen Formen der Erhebungen beschrieben und einige Fragen zur Wahl der Stichprobe erörtert.

Ausführliche Einzelbefragungen

Diese Methode kann in allen Arten und allen Phasen der Evaluation verwendet werden. Es gibt verschiedene Arten von Befragungen, die sich durch Struktur und Zweck unterscheiden: die Befragung durch informelles Gespräch, die halbstrukturierte/auf einem Leitfaden basierende Befragung und die strukturierte Befragung.

Ziel ist es, überwiegend qualitative Informationen von Personen zu sammeln. Qualitative Befragungen arbeiten mit offenen Fragen, die den Befragten viel Spielraum beim Antworten lassen. Es werden vor allem Eindrücke und Deutungen der Interviewten als Informationen berücksichtigt. Im Gegensatz hierzu werden in quantitativen Befragungen geschlossene Fragen gestellt. Damit sind Fragetypen gemeint, für die es vorformulierte Antwortkategorien gibt, wie »Ja/Nein«, Stufen für den Grad einer Zustimmung oder Ablehnung oder die Auswahl von verschiedenen Aussagen.

Qualitative bzw. quantitative Informationen können außerordentlich vielfältig sein: sie reichen von spontanen Äußerungen bis zur wohlüberlegten Beurteilung, Rangreihenfolge, Präferenz, Meinung, Stimmungsbilder, Werte, Faktenwissen, Handlungswissen, Berichte über vergangene Handlungen und zukünftige Verhaltensabsichten.

Auch die zu befragenden Personengruppen können sich deutlich hinsichtlich ihrer Perspektive und möglichen Antworten unterscheiden: Sind die Personen von der Maßnahme betroffen, z. B. als Anwohner bei der Einführung einer neuen Parkraumbewirtschaftung oder profitieren sie davon z. B. als Gäste eines Veranstaltungshauses von der neuen Straßenbahnanbindung? Sind sie an der Planung und Umsetzung der Maßnahme beteiligt, z. B. als Angestellte der Verwaltung oder des Bauunternehmens? Sind sie Vertreterinnen und Vertreter der Öffentlichkeit wie Journalisten, Politiker, Verbraucherorganisationen? In jedem Fall werden alle Personen oder Gruppen ihre jeweils eigenen Sichtweisen und Kompetenzen in das Projekt einbringen können. Als Faustregel kann gelten, dass die Fragen umso offener gestellt werden sollten, je weniger über die unterschiedlichen Perspektiven im Vorfeld bekannt ist. Nur so haben die Befragten die Möglichkeit, in ihren eigenen Worten ihr Erleben oder Verhalten zu schildern und damit das Projekt um neue Sichtweisen zu bereichern.

Einzelbefragungen wurden auch im Projekt Schulumgebung in Evalstadt angewendet. Die Evaluatoren und Evaluatorinnen trafen sich für 30–45 Minuten mit verschiedenen Lehrkräften aus den teilnehmenden Schulen und führten eine teilstrukturierte Befragung durch. Alle Gespräche wurden aufgezeichnet und später transkribiert. Die anschließende Textanalyse zeigte, dass die Eltern, die nicht von Anfang an eingebunden worden waren, auch später schwer erreicht werden konnten. Außerdem zeigte die Analyse, dass das Programm dort erfolgreicher war, wo die beteiligten Lehrkräfte oder die Schulleitung während des laufenden Programmes nicht gewechselt hatten. Beide Erkenntnisse wurden erst durch die Befragung und den Vergleich zwischen den Schulen deutlich.

Im Vergleich zu anderen Methoden ist der Aufwand für die Vorbereitung, die Durchführung, die Dateneingabe durch Abschrift der Interviews und die Auswertung mit Hilfe einer strukturierten Inhaltsanalyse höher. Das wirkt sich meistens auch auf die Kosten aus. Ein ebenfalls nicht zu unterschätzender Effekt ist ein möglicher Einfluss der Interviewerperson auf die Befragten. Diese Forschungsmethode ist nicht für die Durchführung mit vielen Interviewpartnern geeignet. Sie ist jedoch sinnvoll bei allen Maßnahmen, die sich mit sehr innovativen oder komplexen Fragestellungen befassen, für die wenige oder keine Sekundärdaten vorliegen und für die wenige Erfahrungswerte aus ähnlichen Maßnahmen vorliegen. Weitere Informationen zur Analyse von Befragungsdaten gibt es in Kapitel 3.

Umfragen mit Fragebogen

Der Einsatz eines schriftlichen Fragebogens gehört zu den bekanntesten Methoden und ist oft die erste Wahl. Die Vorteile sind, dass auf diese Weise auch eine große Gruppe von Personen befragt werden kann (großer Stichprobenumfang) und dass alle exakt die gleichen Fragen in der gleichen Reihenfolge gestellt bekommen. Die Antworten sind verhältnismäßig einfach auszuwerten. Die Antworten zwischen den Teilnehmenden einer Umfrage können gut miteinander verglichen werden. Eine Fragebogenumfrage kann persönlich durchgeführt werden oder auch schriftlich per Post, E-Mail, über Internetseiten oder mündlich über das Telefon. Das Verteilen und wieder Einsammeln der Bögen kann oft leicht organisiert werden, z. B. durch das Auslegen in Informationszentren oder Verteilen in einem Bus.

Ein Fragebogen kann offene und geschlossene Fragen beinhalten. Je standardisierter ein Fragebogen ist, desto mehr geschlossene Fragen enthält er.

Bei einer offenen Frage ist die Beantwortung frei, d. h. der Befragte schreibt in eigenen Worten die Antwort in das vorgesehene Feld. Eine typische Fragestellung wäre: »Was gefällt Ihnen besonders gut an …?«. Auf diese Weise werden qualitative Daten gesammelt, die entsprechend kategorisiert und ausgewertet werden müssen (siehe Kapitel 3.3 auf Seite 45). Diese offenen Fragen eignen sich immer dann, wenn die Bandbreite der möglichen Antworten sehr individuell ist, z. B. also bei Fragen nach persönlichen Eindrücken.

Geschlossene oder standardisierte Fragen geben mögliche Antworten oder Antwortkategorien vor und werden oft durch Ankreuzen beantwortet. Eine typische Frage wäre: »Kennen Sie das neue Tageticket?« und mögliche Antwortkategorien wären: »ja«, »nein«, »keine Angabe«. Eine Mischung aus standardisierten und offenen Fragen bietet sich in der Praxis oft an und so können die Vorteile beider Fragearten genutzt werden (vergleiche auch Abschnitt zu Skalen weiter hinten).

Bevor ein eigens für die Maßnahme entwickelter Fragebogen breit verteilt wird, sollte er in jedem Fall in einem Vortest erprobt werden. Dazu wird der Fragebogen an eine kleine Stichprobe verteilt und anschließend ausgewertet. So können eventuell missverständliche Fragen oder fehlende Antwortkategorien überarbeitet werden.

Kasten 5: Codierung der Fragebogen-ID bei abhängigen Stichproben

Fragebögen müssen aus Gründen der Datensicherheit und des Datenschutzes anonym bleiben. Wenn für ein Paneldesign die gleichen Personen zwei- oder mehrfach befragt werden sollen, muss es aber möglich sein, die verschiedenen Fragebögen in der Analyse der gleichen Person zuzuordnen.

Daher bittet man die Befragten, vor dem Ausfüllen um die Erstellung eines individuellen Codes, der die Anonymität wahrt, aber nicht vergessen werden kann. Für einen sechsstelligen Code empfehlen wir folgende Fragen:

1. Erster Buchstabe des Geburtsnamens der Mutter
2. Letzter Buchstabe des eigenen Geburtsnamens
3. Tag und Monat der Geburt (Form: ttmm)

Auf diese Weise erhalten Sie von einer Person mit dem Namen »Schmidt«, geboren am 20. September, dessen Mutter eine geborene »Maier« ist, folgende ID: MT2009.

1. Was ist das Ziel Ihrer Reise?

- Arbeitsplatz
- Schule
- Freizeit
- Arztbesuch
- Zuhause
- anderes Ziel

2. Wie oft benutzen Sie die Buslinie?

- 5-7 Mal pro Woche
- 3-4 Mal pro Woche
- 1-2 Mal pro Woche
- 1-3 Mal pro Monat
- seltener

3. Wie zufrieden sind Sie mit den folgenden Elementen der Busse?

	sehr unzufrieden	unzufrieden	weder-noch	zufrieden	sehr zufrieden	Weiß nicht
Belüftung	☐	☐	☐	☐	☐	☐
Sitzkomfort	☐	☐	☐	☐	☐	☐

Abbildung 2-10: Auszug aus einem Fragebogen, der in einem Bus verteilt wurde

Fokusgruppen

Fokusgruppen sind moderierte Gruppendiskussionen mit einem bestimmten Fokus. Sie werden dann eingesetzt, wenn es darum geht, Vertreterinnen und Vertreter auch unterschiedlicher Interessen miteinander ins Gespräch zu bringen. Dies kann besonders nützlich für Bereiche oder Maßnahmen sein, die komplex sind, oder bei denen es wichtig ist, unterschiedlichen Anforderungen verschiedener Nutzergruppen gerecht zu werden.

In Fokusgruppen wird die Interaktion, Kreativität und Spontanität der Teilnehmenden genutzt. Die Moderatorin oder der Moderator stellt die Themen vor und lenkt die Diskussion durch spezifische Fragen. Er oder sie fasst Ergebnisse der Diskussion zusammen und dokumentiert gegebenenfalls verbleibende Meinungsunterschiede. Fokusgruppen sind nicht automatisch konfliktgeladene Streitgespräche. Ganz im Gegenteil können gerade die unterschiedlichen Perspektiven zu einem ganzheitlichen Ansatz führen. Beispielsweise können Fokusgruppendiskussionen dabei helfen, bei der Planung der Innenraumgestaltung von Straßenbahnen die Bedürfnisse aller Nutzergruppen angemessen zu berücksichtigen und mögliche Konflikte schon im Vorfeld zu entschärfen.

Fokusgruppen können in jeder Phase einer Evaluation eingesetzt werden. Vor dem Einsatz von Fokusgruppen ist es wichtig, dass die relevanten Zielgruppen bestimmt werden. Je nach Fragestellung kann dann abgewogen werden, ob es sinnvoller ist, Fokusgruppendiskussionen mit gemischten Zielgruppen durchzuführen oder lieber getrennt voneinander. Im obigen Beispiel der Straßenbahngestaltung ergibt es mehr Sinn Sprecherinnen und Sprecher für die Gruppen »Fahrrad«, »Rollstuhl« »Kinderwagen«, »ältere Menschen«, »Berufspendelnde«, »Reisende mit Gepäck« und »Jugendliche« gleichzeitig an der Planung zu beteiligen. Die Moderierenden sind dann aufgefordert, alle Gruppen und Personen zu Wort kommen zu lassen und den Kompromiss neutral zu verhandeln. Durch eine gut moderierte Diskussion zwischen allen kann auch das gegenseitige Verständnis unterstützt, eine breite Zustimmung oder ein tragfähiger Kompromiss erreicht werden.

In Projekten, in denen die Diskussionsteilnehmenden einen direkten oder persönlichen Bezug haben oder gar negativ Betroffene der Maßnahme sind, kann es passieren, dass diese in der Fokusgruppe vor allem bereits bekannte Argumente austauschen und sich die Diskussion schnell im Kreis dreht oder hitzig wird – Sie also keine neuen Informationen sammeln können. Befragen Sie die Gruppen getrennt voneinander, werden sich die Teilnehmenden vermutlich anders verhalten, offener sprechen, die Fragestellung kreativer von mehreren Seiten durchdenken. Daher ist die Fokusgruppe eine gute Evaluierungs- und auch Planungstechnik für alle innovativen Pilotmaßnahmen.

Beobachtungstechniken

Bei einer Beobachtung werden Menschen während ihres tatsächlichen Verhaltens betrachtet. Die geschulten Beobachterinnen und Beobachter sind entweder sichtbar oder nicht sichtbar für die Beobachteten. Stellen Sie sich die Haltestelle einer S-Bahn vor: Wie verhalten sich Menschen auf dem Bahnsteig? Wie viele sind alleine, wie viele in Gruppen unterwegs? Welchen Abstand halten sie zu Wänden oder Objekten? Wie schnell gehen sie? Warten sie in Gruppen, in einer Schlange oder halten sie dabei einen möglichst großen Abstand zu anderen? Stehen sie alle in einem Abschnitt oder verteilen sie sich über das gesamte Gleis? Sind sie beim Ein- und Aussteigen eher diszipliniert oder eher chaotisch? Was passiert, wenn es anfängt zu regnen?

Die direkte Beobachtung hat den Vorteil, dass der gesamte Kontext erlebbar wird. Wie ist die Geräuschkulisse auf dem Bahnsteig? Wie riecht es auf der Treppe? Wie glatt ist der Boden? Wie lenkt die Beleuchtung? Welche Hinweise oder Schilder können wahrgenommen werden? Wie verständlich sind Durchsagen? Ausgebildete Evaluatorinnen und Evaluatoren können auch Phänomene wahrnehmen, die der Aufmerksamkeit der anderen entgehen oder durch eine Befragung nicht erfasst werden können.

Eine Variante ist der Einsatz von Blickbewegungskameras. Diese können wie eine Brille getragen werden und zeigen bei der Auswertung der Filme exakt die Perspektive der Nutzer. Diese aufwändige Methode ist besonders hilfreich in allen sicherheitsrelevanten Bereichen, z. B. der Planung von Notaus-

gängen oder aber um die spezifischen Bedürfnisse von Nutzern kennenzulernen, z. B. von Menschen mit eingeschränkter Mobilität. Kombiniert mit der Technik des »lauten Denkens« kann man so enorm viel über die Situation lernen.

Mit diesen Techniken ist es möglich, ein vielseitiges Bild der zu evaluierenden Situation zu gewinnen, die weit über den Einsatz von Fragebögen und Befragungen hinausgeht. In der Kombination mit anderen Methoden ist sie daher besonders stark.



Außer selbst zum Ort der Maßnahme zu gehen und mehr oder weniger standardisierte Berichte und Beobachtungsprotokolle zu erstellen, wird immer häufiger auch die Videoüberwachung genutzt. Dies kann eine effiziente Methode sein, hat aber manchmal den Nachteil, dass sie nur einen Teil der Situation erfassen kann. Viele eventuell wichtige Rahmenbedingungen, wie Beleuchtung, Einsehbarkeit, Übersichtlichkeit, Gerüche, Bodenbeschaffenheit etc. können nicht über das Videobild übertragen werden.

Vor- und Nachteile der Methoden zur Datenerhebung

Die Auswahl einer geeigneten Methode erfordert eine sorgfältige Prüfung vieler Faktoren. Neben Fragen zu Stichprobengröße, Zielgruppen, Kosten der Umfrage, erwartete Antwortquoten und Datenqualität können auch länderspezifische Datenschutzbestimmungen die Entscheidung beeinflussen. Dieser Abschnitt enthält eine Zusammenfassung der gängigen Methoden zur Datenerhebung.

Die persönliche Datenerhebung liefert in der Regel die bestmögliche Abdeckung des Themas, erreicht hohe Antwortquoten und liefert die beste Datenqualität. Es überrascht aber nicht, dass persönliche Befragungen auch zu den teuersten Methoden gehören. Der größte Aufwand besteht oft darin, die Zielgruppen zu erreichen und zur Teilnahme zu bewegen. Aus diesem Grund werden häufiger Telefon- und Postverfahren verwendet, auch wenn dies eine Kompromisslösung darstellt.

Bei einer Telefonbefragung werden die zu Befragenden angerufen und telefonisch interviewt. Meistens werden diese Interviews auch aufgezeichnet und später ausgewertet. Je nach Art und Umfang der Befragung kann die Vorbereitung wie der Kauf der Telefonnummern und die Einrichtung der entsprechenden Systeme zum Aufzeichnen aufwändig sein. Hier sollte man ggf. mit einem, auf diese Methode spezialisierten Unternehmen zusammenarbeiten. Im Vergleich zum Postversand haben Telefonbefragungen eine höhere Antwortquote und sind daher repräsentativer, sofern mögliche Abweichungen im Stichprobenrahmen der Telefonnummern berücksichtigt werden.

Der Hauptvorteil von schriftlichen Fragebögen ist, dass sie preiswert sind und geografisch großflächig verteilt werden können. Das Versenden und Zurückschicken von Fragebögen nimmt jedoch viel Zeit in Anspruch. Bei diesen Methoden sind geringe Antwortraten und unvollständig ausgefüllte Fragebögen häufige Probleme.

Kasten 6: »Hallo Welt!«

In den letzten Jahren wurde die Verwendung des Internets für Umfragen sehr beliebt. Dies ist häufig einfacher und schneller als die »herkömmlichen« Papierfragebögen. Es gibt jedoch einige Aspekte, die bei der Nutzung des Internets zu berücksichtigen sind. Verteilung und Zugang sind zwischen Ländern und Städten unterschiedlich. Man kann davon ausgehen, dass in Island 100% der Bevölkerung Internet-Zugang haben, in Italien sind es aber nur etwa 50%. Daher ist im Einzelfall eine echte Zufallsstichprobe schwer zu erreichen, die Ergebnisse können durch die Zugänglichkeit (oder Nicht-Zugänglichkeit) zum Internet bereits verfälscht werden. Gleiches gilt auch für Umfragen über social media. Wenn Sie z. B. eine Befragung über facebook starten, erreichen Sie nur die facebook-Nutzenden, und nicht diejenigen, die dort nicht angemeldet sind.

Andererseits gibt es auch wesentliche Vorteile. Die Ergebnisse gelangen normalerweise direkt in eine Datenbank, wodurch auf die Antworten sofort zugegriffen werden kann. Die Umfragen sind auch relativ günstig durchzuführen, da es bereits zahlreiche Plattformen für die Durchführung und Präsentation von Online-Fragebögen gibt, z. B. survey monkey, esurveys.pro, NetQ und QuestionPro. Die Umfragen können auf die spezielle Anwendung angepasst werden, wodurch allgemeine Aussagen wie beispielsweise »wenn nicht, weiter mit Nummer …« vermieden werden können.

Skalen

Es gibt vier verschiedene Arten von Skalen, die im Folgenden kurz vorgestellt werden. Je nach Skala sind in der späteren Auswertung unterschiedliche Rechenoperationen oder statistische Verfahren zulässig anwendbar.

Die Nominalskala ermöglicht eine Unterscheidung zwischen Gleichheit und/oder Verschiedenheit. Mit der Ordinalskala können Objekte geordnet werden. Bei einer Intervallskala sind die Unterschiede zwischen zwei aufeinanderfolgenden Objekten gleich groß. Die Verhältnisskala enthält einen »natürlichen Nullpunkt«. Der gemessene Wert 0 zeigt die aktuelle Abwesenheit der gemessenen Eigenschaft.

Skalenart	Ermöglicht Aussagen zu	Beispiele
Nominal-skala	Gleichheit Ungleichheit	Geschlecht (z. B. männlich, weiblich); Wahl des Verkehrsmittels (z. B. zu Fuß, Fahrrad, Auto, öffentliche Verkehrsmittel usw.)
Ordinal-skala	größer/kleiner als Beziehungen	Kundenzufriedenheit (z. B. sehr zufrieden, zufrieden, weder noch, unzufrieden, sehr unzufrieden)
Intervall-skala	Gleichheit der Unter-schiede	Jahreszahlen, Geburtsjahr Startzeit, Zielzeit
Verhältnis-skala	Gleichheit der Verhält-nisse, Nullpunkt	Geschwindigkeit (z. B. 30km/h), Entfernung (Wegstrecke in Metern oder Kilometern)

Tabelle 2-3: Beispiele für die vier Arten von Skalen

Kasten 7: Wahl der richtigen Skala

Das menschliche Urteilsvermögen ist meistens subjektiv. Um Meinungen möglichst objektiv und vergleichbar zu erfassen, bieten sich unter anderem sogenannte Ratingskalen an (Rating engl. = Einschätzung).

Es wird dabei angenommen, dass die einzelnen Stufen der Bewertung die gleichen Abstände haben, also intervallskaliert sind, und sich das Merkmal auf einem Kontinuum beschreiben lässt. Die Bewertung kann numerisch (quantitativ) oder qualitativ sein.

Eine unipolare Skala fordert einen Befragten z. B. auf, ein tatsächliches Verhalten zu beschreiben. Auf die Frage: »Wie häufig nutzen Sie die Stadtbusse?« könnten mögliche Antwortkategorien »5–7 Tage pro Woche«, »3–4 Tage pro Woche«, »1–2 Tage pro Woche«, »1–3Tage im Monat«, »seltener« sein. Eine bipolare Skala fordert einen Befragten auf, zwischen zwei gegensätzlichen Eigenschaften zu wählen und den relativen Anteil dieser Eigenschaften festzulegen. Eine bekannte bipolare Skala ist die Likert-Skala. Die Form einer typischen Likert-Skala mit fünf Stufen ist: »trifft nicht zu«, »trifft eher nicht zu«, »weder noch«, »trifft eher zu«, »trifft zu«. Eine Skala mit nur vier Optionen verzichtet auf die mittlere Option »weder noch«. Dies bringt die Antwortenden dazu, sich für eine Richtung zu entscheiden. Die folgenden Tipps helfen beim Erstellen eigener Skalen:

Nicht mehr als sieben Werte auf einer Skala einsetzen: Studien zeigen, dass Menschen kaum in der Lage sind, ihren Standpunkt auf einer Skala von mehr als sieben Möglichkeiten einzutragen. Es gibt aber keine perfekte Anzahl, sie hängt von der Frage ab. Am häufigsten genutzt werden Fünf-, Vier- und Drei-Punkt-Skalen.

Nummerierte Skalen bringen ungenauere Ergebnisse: Skalen, die nur numerisch gekennzeichnet sind, z. B. »1« bis »5« bringen ungenauere Ergebnisse als Skalen mit einer verbalen Kennzeichnung wie z. B. »gut« oder »schlecht«. Bei Verwendung nummerierter Skalen werden zusätzliche verbale Erläuterungen empfohlen, z. B. »sehr gut« über der 1, »zufriedenstellend« über der 3 und »mangelhaft« über der 5.

Größe und Auswahl der Stichproben

Es ist wichtig, die Größe der benötigten Stichprobe bereits im Vorfeld sorgfältig zu bestimmen: Je mehr ausgefüllte Fragebögen Sie erhalten, desto größer ist zwar die Sicherheit, aber desto höher sind die Kosten. Daher gilt es, maximale Sicherheit zu minimalen Kosten zu erhalten. Die statistische Sicherheit wird von vier Faktoren bestimmt, die im Folgenden vorgestellt werden:

Größe der Stichprobe: Für eine Umfrage ist die Grundgesamtheit aller Personen, über die eine Aussage getroffen werden soll, von Bedeutung, das ist die so genannte »Population«. Da in aller Regel nicht alle Personen befragt werden können oder sollen, gilt es, aus dieser Gesamtgruppe eine Stichprobe auszuwählen, die quasi stellvertretend für alle antworten. Die für die Fragestellung optimale Größe der Stichprobe kann mathematisch berechnet werden. Faustregeln für die Schätzung von Zufallsstichproben werden in Kasten 8 vorgestellt. Da erfahrungsgemäß nicht alle Personen auf eine Befragung antworten, die so genannte Antwortquote also nie 100% beträgt, müssen entsprechend mehr Personen befragt werden, damit die erforderliche Stichprobengröße erreicht wird. Das folgende Beispiel spielt dies durch:

In Evalstadt gibt es ein Parkraummanagementsystem. Im Stadtteil A gibt es 4.000 Kraftfahrzeughalter, das ist die Population des Stadtteils. Für eine Umfrage zu der Funktionalität einer neuen Online-App soll die Netto-Stichprobengröße 400 betragen. Dazu werden aus den 4.000 Haltern per Zufall 700 (die so genannte Bruttostichprobe) zur Teilnahme an einer Online-Umfrage ausgesucht und angeschrieben. 420 füllen den Fragebogen aus, damit beträgt die Antwortquote 60% und die Netto-Stichprobengröße ist erreicht. Generell ist eine möglichst hohe Antwortquote anzustreben, denn sie bildet die Grundlage, auf der die Berechnungen durchgeführt und damit die Rückschlüsse auf die Population gezogen werden. Oft werden bei schriftlichen Umfragen aber nur geringere

Rückläufe erreicht, im Schnitt zwischen 15–30%. Dies ist bei der Auswahl der Brutto-Stichprobe zu beachten oder alternativ durch eine Methode mit einer höheren Rücklaufquote auszugleichen, wie beispielsweise durch die persönliche Befragung.

Es stehen eine Reihe von Online-Stichprobengrößenrechnern zur Verfügung (siehe Kapitel 9.3 auf Seite 77).

Gewünschte Segmentierung der Analyse: In der Regel wird ein Datensatz hinsichtlich einzelner Fragestellungen als Ganzes ausgewertet, in Bezug auf spezifische Fragen auch in Segmenten. Im Beispiel zu der Funktionalität einer neuen Online-App könnte es sinnvoll sein, das Antwortverhalten nach Altersgruppen getrennt auszuwerten, da ältere PKW-Halter internetfähige Anwendungen eventuell weniger häufig oder anders nutzen als jüngere. Jede Altersgruppe (Segment) stellt dann eine eigene Population dar. Werden Schlussfolgerungen aus der Analyse eines Segments gezogen, muss die statistische Sicherheit auf dieses Segment bezogen sein, nicht auf die Gesamtpopulation.

Grad der Varianz bei den Antworten der Population: Die Stichprobengröße hängt auch davon ab, wie groß das Spektrum der möglichen Antworten auf die Frage ist. Je größer das Spektrum ist, desto größer muss auch die Stichprobe sein – und umgekehrt. Bei den meisten Befragungen ist im Vorfeld wenig darüber bekannt, wie groß diese Varianz ist, daher sollte die Stichprobengröße besser höher angesetzt werden.

Fehlertoleranz: Da nur selten die komplette Grundgesamtheit befragt wird, im Beispiel alle Kraftfahrzeughalter in Evalstadt, können die durch eine Stichprobe ermittelten Werte immer nur Näherungen sein. Es gibt also eine Abweichung vom tatsächlichen Wert. Wie groß diese Abweichung sein darf, d. h. wie groß dieser so genannte Fehler ist, kann festgelegt werden. In aller Regel liegt er zwischen 5% und 10%. Je größer die Investitionsentscheidung oder je bedeutender die Auswirkungen einer Maßnahme ist, desto niedriger wird die Fehlertoleranz sein. Der anerkannte Standard für die statistische Sicherheit, das so genannte Konfidenzintervall, liegt zwischen 90% und 95%.

Kasten 8: Wie viele sind genug

Als grobe Faustregel gilt, dass 200 Antworten für die meisten Annahmen und Parameter eines Umfrageprojekts eine recht gute Genauigkeit der Umfrage ergeben – außer bei der Analyse innerhalb einzelner Segmente. Für eine gerade noch akzeptable Genauigkeit werden rund 100 Antworten benötigt.

Die oft als ausreichend angenommene Größe von 30 Antworten bieten nur dann eine annehmbare Genauigkeit, wenn es sich um entweder eine sehr kleine Population handelt, das Spektrum der möglichen Antworten sehr gering ist oder es zunächst nur um eine sehr ungenaue Einschätzung geht.

Letztendlich lassen sich keine allgemeingültige Aussagen über den Umfang einer Stichprobe treffen. Für die Evaluation eines sehr wenig streuenden Merkmals kann ein relativ kleiner Stichprobenumfang ausreichen. Ist die Variation eines Merkmals sehr groß, muss ein größerer Stichprobenumfang realisiert werden. Wenn es um Differenzen von zwei Gruppen geht, können schon kleine Stichprobenumfänge ausreichen. Sind jedoch die Unterschiede in den Kennwerte der zu vergleichenden Gruppen gering, braucht es wiederum größere Stichproben.

Ausgehend von dem Evaluationsziel bzw. der Forschungsfrage und der Recherche über die Grundgesamtheit und den Ergebnissen von Pretests muss abgewogen werden, welcher Stichprobenumfang ausreichend ist. Auch können Stichprobenumfänge gesenkt werden, wenn eine adäquate geschichtete Stichprobe, z. B. nach soziodemographischen Merkmalen, vorgenommen werden.



Grundlegende Anregungen für die Durchführung einer Umfrage

Folgende Ratschläge helfen bei der Planung einer Umfrage:

- Sammeln Sie nur relevante Daten. Vermeiden Sie das Erheben überflüssiger Informationen. Dies treibt nicht nur den Aufwand der Auswertung in die Höhe, sondern kann auch einen großen Einfluss auf das Antwortverhalten haben. Im schlimmsten Fall riskieren Sie, dass die Befragten abbrechen.
- Planen Sie den nötigen Grad der Detailliertheit und der Genauigkeit vor der Datenerhebung.
- Berücksichtigen Sie, dass auch die Verwaltung der erhobenen Daten eine komplexe Aufgabe ist und Erfahrung voraussetzt. Stimmt die Qualität und der Umfang der erhobenen Daten? Werden sie gemäß den Richtlinien für Datenschutz gespeichert und gesichert?
- Nennen Sie gegenüber den Befragten den Grund und die Ziele der Umfrage. Wer ist der Auftraggeber, was geschieht mit den Informationen, wozu dienen sie?
- Führen Sie in jedem Fall einen Vortest (Pretest) durch.

3 Dateneingabe und Datenauswertung

Zur »richtigen« Analyse von Daten wurden schon viele Bücher geschrieben. Die folgenden Unterkapitel geben eine knappe Übersicht. Kapitel 3.1 und 3.2 beschreiben die quantitative Datenanalyse, während das Kapitel 3.3 eine Einführung in die qualitative Datenanalyse gibt.

3.1 Datenaufbereitung und Datensicherung

Mit Hilfe der in den vorangegangenen Abschnitten vorgestellten Methoden werden viele Daten erhoben. Um diese so zu strukturieren, dass sie in einer Tabelle dargestellt und anschließend ausgewertet werden können, ist ein so genannter Codeplan erforderlich. Dies ist eine Liste aller Variablen mit allen Antwortmöglichkeiten. Darin wird jeder Antwort ein bestimmter Zahlenwert (Code) zugewiesen. Die gesammelten Daten werden dann nach diesem Plan codiert.

Besitzen Sie ein Auto?	Ja (⇢ codiert mit 1) Nein (⇢ codiert mit 2)	Nominalskala
Welche Strecke sind Sie heute mit dem Auto gefahren?	0 km (⇢ codiert mit 1) bis 5 km (⇢ codiert mit 2) ›5 ‹10 km (⇢ codiert mit 3) …	Verhältnisskala
In welchem Monat endet Ihre Jahreskarte?	Januar (⇢ codiert mit 1) Februar (⇢ codiert mit 2) … Dezember (⇢ codiert mit 12)	Intervallskala
Wie bewerten Sie den Stadtbus hinsichtlich des Komforts der Sitze?	gar nicht komfortabel (⇢ 1) wenig komfortabel (⇢ 2) etwas komfortabel (⇢ 3) sehr komfortabel (⇢ 4)	Ordinalskala

Abbildung 3-1: Beispiel Codeplan (mittlere Spalte)

Vor der Dateneingabe sollte jedem Fragebogen, bzw. jedem Datenblatt der befragten Person eine Identifizierungsnummer (ID) zugeordnet werden. Am besten werden dazu die Bögen durchnummeriert. Auf diese Weise können ausgefüllte Fragebögen später erneut mit der Dateneingabe verglichen werden, z. B. um mögliche Fehler bei der Dateneingabe zu korrigieren.

Den Antwortmöglichkeiten sind laut Codeplan Zahlenwerte zugeordnet (Abb. 3-1 mittlere Spalte), diese werden an Stelle der Antwort in die Tabelle eingetragen. Für jede befragte Person wird eine neue Zeile angelegt. Jede Spalte steht für eine Frage.

ID	Autobesitz	Strecke Auto	Ende Jahreskarte	Sitzkomfort
…				
39	2	1	3	3
40	2	2	3	3
41	1	3	12	3
42	1	3	10	4
43	2	1	9	3

Tabelle 3-1: Beispiel Datentabelle

Auf diese Weise werden die Daten z. B. in die Tabellenkalkulationssoftware (Microsoft Excel) oder eine Statistiksoftware (z. B. SPSS, SAS, Statistica usw.) per Hand eingegeben. Diese Art der Dateneingabe wird für eine kleine Zahl von Fällen und bei relativ kurzen Fragebögen empfohlen (bis zu 300 Fälle mit 30 Variablen). Der Prozess der Dateneingabe ist zeitaufwändig, arbeitsintensiv und leider auch fehleranfällig. Das liegt daran, dass die Dateneingabe hochkonzentriert erfolgen muss. Aber: Finanzielle und technische Einschränkungen bei vielen Projekten lassen keine andere Wahl, als die Daten z. B. in Microsoft Excel einzugeben. In diesen Fällen ist es oft sinnvoll, die Dateneingabe in einem Team von zwei Personen durchzuführen. Eine Person liest die Zahlen laut vor und die andere gibt diese über die numerische Tastatur ein. Alternativ können stichprobenartig Kontrollen durch eine weitere Person durchgeführt werden.

Im Beispiel (Abbildung 3-1) wird davon ausgegangen, das die Befragung mit Hilfe eines schriftlichen Fragebogens (»paper & pencil«) und anschließend die Dateneingabe per Hand vorgenommen wurde. Dies ist sicherlich immer noch die am weitesten verbreitete Durchführung, doch wird sich das mit der zunehmenden Verfügbarkeit von preiswerten Computern und Tablets, der weiteren Vereinfachung von Onlinefragebögen, sowie vor allem der zur Verfügung stehenden Software ändern. Die direkte Dateneingabe durch die Befragten selbst spart eine Menge Zeit. Sie kann auf unterschiedliche Weise erfolgen. Idealerweise gibt es eine programmierte Eingabemaske (über Software wie Data Entry oder Microsoft Access), die genau wie der verwendete Fragebogen aussieht. Bei Verwendung einer Eingabemaske können auch nicht plausible Eingaben (z. B. Alter: 200) durch entsprechende Fehlerprüfungen verhindert werden. Für sehr umfangreiche Fragebögen oder bei großen Stichproben ist die Programmierung einer Eingabemaske sehr empfehlenswert. Eine Eingabemaske kann auch automatische Filter enthalten, welche die dateneingebende Person je nach Antwort durch verschiedene Abschnitte eines Fragebogens führt. Zu beachten sind der zusätzliche Zeitaufwand für die Programmierung der Eingabemaske sowie die Anschaffungskosten für die Hard- und Software.

Eine andere Möglichkeit der automatisierten Dateneingabe ergibt sich durch Scanner mit entsprechender Software. Diese Methode spart ebenfalls viel Zeit. Sie ist sinnvoll bei einer hohen Zahl von Fällen, bei denen die Fragebögen nur wenige Seiten haben. Die Fragebögen müssen aber ordnungsgemäß und sehr zuverlässig ausgefüllt werden. Sitzen die Kreuze der Antworten nicht exakt am richtigen Punkt oder sind zu schwach, kann es Probleme beim Scanvorgang geben, dann erhöht sich der Aufwand für die manuelle Nachbearbeitung. In diesem Fall muss die Dateneingabe letztendlich einzeln erfolgen. Ein weiterer Nachteil ist, dass der Entwurf des Fragebogens systemkompatibel zur Scannersoftware sein muss. Dadurch werden die Gestaltungsmöglichkeiten durch die Software beschränkt und die entstandenen Fragebögen wirken möglicherweise nicht sehr ansprechend. Ein schön und vor allem übersichtlich gestalteter Fragebogen ist mitentscheidend für eine hohe Antwortquote und unterstützt auch eine korrekte Beantwortung.

	Paper & Pencil Eingabe per Hand	Direkte Eingabe Online oder über Tablet	Scannen der Fragebögen mit entsprechender Software
Vorteile	· geringe Vorarbeiten · geringe technische Voraussetzungen (Computer, Excel)	· geringer Zeit- und Arbeitsaufwand · automatische Fehlerprüfung · Filtereinstellungen	· mittlerer Zeit- und Arbeitsaufwand
Nachteile	· hoher Zeit- und Arbeitsaufwand · Fehler bei der Dateneingabe	· Vorarbeiten bei der Programmierung · Erstellen einer Eingabemaske · Anschaffungskosten für Hard- und Software	· Abhängigkeit von der Hardware · Gestaltung der Fragebögen muss systemkompatibel sein · Fragebögen müssen systemkompatibel ausgefüllt werden
Empfohlene Anwendung	· kleine Anzahl von Fällen · Kurze Fragebögen	· Große Anzahl von Fällen · Umfangreiche Fragebögen	· Große Anzahl von Fällen · Wenige Seiten · Wenn Fragebögen ordentlich und zuverlässig ausgefüllt wurden

Tabelle 3-2: Vergleich verschiedener Methoden zur Dateneingabe

Bei der Dateneingabe empfiehlt es sich, ein paar Regeln zu beachten, damit es nicht zu unbeabsichtigten Manipulationen oder Verfälschungen kommt. Es muss alles genau so eingegeben werden, wie es im Fragebogen angegeben ist, es sei denn, es sind offensichtliche Fehler. Wird etwas anderes als die

Antwort im Fragebogen eingegeben, muss dies dokumentiert werden (z. B. auf dem Fragebogen oder besser in einer Kommentarspalte im Datenblatt). Fehlen viele Antworten, oder wurden die Antworten nicht ernsthaft gegeben, sollte der Fragebogen bei der späteren Analyse und Interpretation gar nicht berücksichtigt werden.

Bei der Dateneingabe können Fehler auftreten. Insbesondere die manuelle Dateneingabe in Tabellenkalkulationssoftware ist häufig fehleranfällig. Es können Tippfehler vorkommen oder es wird versehentlich ein nicht dokumentierter Code eingegeben (d. h. eine Nummer, die nicht im Codeplan angegeben ist).

Aufgrund der Fehlermöglichkeiten bei der Dateneingabe sollte vor dem Auswerten eine Kontrolle auf Vollständigkeit und Ausreißerwerte stattfinden. Eine einfache Kreuztabelle für den ersten Überblick kann sehr hilfreich sein. Ein fünf Jahre altes Kind mit Führerschein oder 10.000 gefahrene Kilometer pro Woche sind nicht plausibel. Wenn die Daten solche offensichtlich fehlerhafte Informationen enthalten, muss dies erneut durch Vergleich von Fragebogen und elektronischem Datensatz geprüft werden. Im besten Fall kann der Fehler korrigiert werden, im schlechtesten Fall muss der Datensatz aus der weiteren Analyse ausgeschlossen werden. Auch solche Fälle sind für die spätere Diskussion der Daten entsprechend zu dokumentieren.



Die erhobenen Daten sind die Grundlage für die Evaluation, Veröffentlichung und Entscheidungsfindung und sind deshalb sehr wertvoll. Sie müssen auf haltbaren und möglichst sicheren Speichermedien gespeichert werden (beispielsweise: USB-Sticks, externe Festplatten usw.). Die Datensicherung muss regelmäßig durchgeführt werden, damit immer die neuesten Daten gesichert sind. Mindestens zwei räumlich getrennt gelagerte Sicherungen sind ratsam. Die Daten müssen regelmäßig auf Vollständigkeit und Unversehrtheit überprüft werden. Und schließlich müssen sie vor unbefugtem Zugriff geschützt werden. Es empfiehlt sich, entsprechende Datenschutzregeln von den beteiligten Personen unterschreiben zu lassen.

3.2 Datenverarbeitung und -analyse

Im folgenden Abschnitt werden die Schritte zur Durchführung statistischer Testverfahren erläutert. Die Abbildung 3-2 zeigt diese im Überblick.

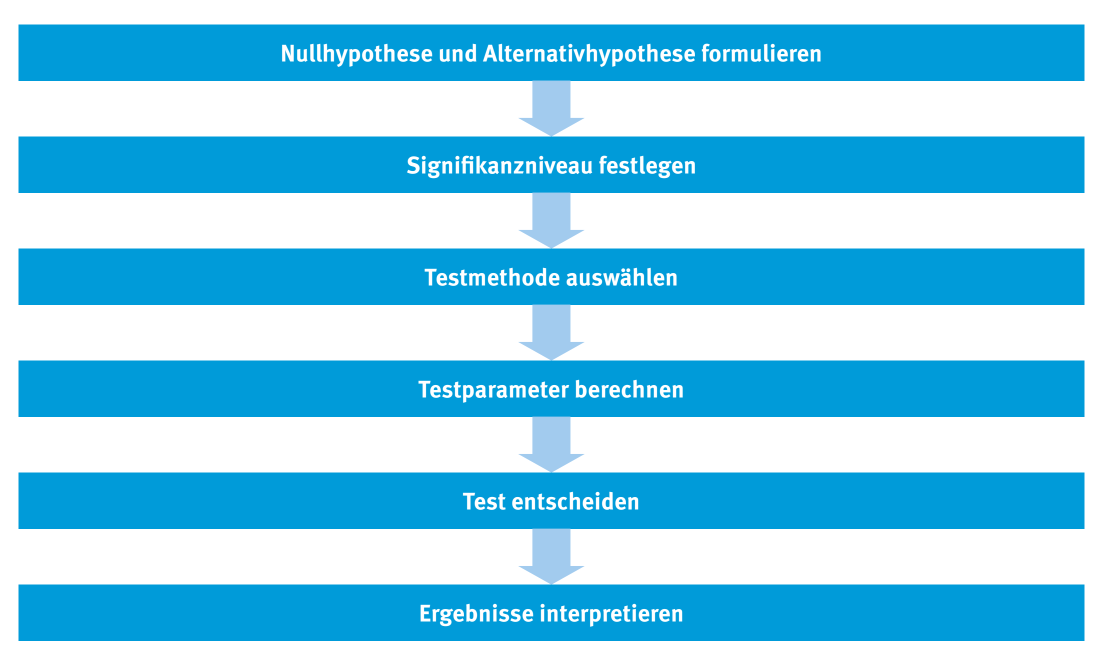


Abbildung 3-2: Durchführungsschritte statistischer Testverfahren

Eine der ersten Aufgaben bei der Datenanalyse bestehen darin, die empirisch erhobenen und ausgewerteten Daten in Tabellenform zusammenzufassen.

Wie oft nutzen Sie die Stadtbusse?	Häufigkeit	Prozent	Kumulierte Prozent
5–7 Tage pro Woche	472	58,1	58,1
3–4 Tage pro Woche	186	2,9	80,9
1–2 Tage pro Woche	86	10,6	91,5
1–3 Tage im Monat	40	4,9	96,4
Seltener	29	3,6	100,0
Summe	**813**	**100,0**	

Tabelle 3-3: Beispieltabelle für eine erste Häufigkeitsanalyse

Der erste Schritt der Deskriptivstatistik oder auch beschreibende Statistik ist die Zählung, in Tabelle 3-3 in der Spalte »Häufigkeit« dargestellt. Sie ist für alle Arten von Skalen möglich. Neben den absoluten Häufigkeiten gibt es kumulative (d. h. fortlaufend aufsummierte Häufigkeiten der Kategorien) und prozentuale Häufigkeiten. Letztere ermöglichen es, Verteilungen mit unterschiedlichen Gesamtausmaßen zu vergleichen. In der Regel sollten diese Häufigkeiten immer mit der entsprechenden Anzahl der befragten Fälle angegeben werden (in der Tabelle als »Summe« bezeichnet). Bei der Auswertung von Skalen müssen häufig Kategorien oder Klassen erstellt werden, um die Eigenschaften der Vertei-

lung besser zu veranschaulichen. Die Anzahl dieser Klassen muss so gewählt werden, dass die Leserinnen und Leser den Überblick behalten.

Statistische Parameter dienen dem Zweck, zusammengefasste Informationen über spezielle Eigenschaften bei der Verteilung von Merkmalen zu liefern. Es gibt zwei Arten von Parametern: Lageparameter und Verteilungsmaße. Welche Parameter berechnet werden können, hängt von der jeweiligen Skala ab.

Lageparameter werden auch als »Maße der zentralen Tendenz« bezeichnet. Sie ermöglichen Aussagen über die Verteilung der Eigenschaften in Bezug auf die Frage, welcher Wert die Gesamtverteilung am besten repräsentiert. Wichtige Lageparameter sind Modalwert, Median und arithmetisches Mittel (siehe Kasten 9).

Kasten 9: Lageparameter (»Maße der zentralen Tendenz«)

Der **Modalwert** oder **Modus** ist der Wert einer Verteilung, der am häufigsten auftritt, d. h. die am häufigsten auftretende Kategorie. Er kann auf jedem Skalenniveau bestimmt werden.

Der **Median** teilt die Häufigkeitsverteilung in zwei gleich große Teile. Bei einer kumulierten Darstellung markiert der 50%-Wert den Median.

Das **arithmetische Mittel** ist die Summe der Messwerte, geteilt durch die Anzahl der Messwerte.

Sind zwei Verteilungen hinsichtlich ihrer zentralen Tendenzen ähnlich, können sie dennoch in Bezug auf die Streuung der einzelnen Werte sehr unterschiedlich sein. Während der Lageparameter also angibt, welcher Wert eine gegebene Verteilung am besten repräsentiert, informiert das Streuungsmaß über die Unterschiede bei den Messwerten. Somit dient das Streuungsmaß einer genaueren Beschreibung der Verteilung und quantifiziert die in der Umfrage erkannten Unterschiede. Spannweite, Varianz und Standardabweichung sind die häufigsten Streuungsmaße (siehe Kasten 10).

Kasten 10: Arten von Streuungsmaßen

Die **Spannweite** (oder Variationsbreite) umfasst den gesamten Bereich der Variationen aller Messwerte. Sie wird als Differenz zwischen dem maximalen und dem minimalen Wert definiert.

Die **Varianz** und die Standardabweichung sind die häufigsten Kennzahlen einer Verteilung. Für ihre Berechnung werden intervallskalierte Daten benötigt.

Die Varianz zeigt den Unterschied zwischen den Werten einer gegebenen Verteilung. Sie wird berechnet, in dem die Summe der quadrierten Abweichungen vom arithmetischen Mittel durch die Anzahl der Messwerte geteilt wird.

Die **Standardabweichung** wird als Quadratwurzel der mittleren quadratischen Abweichung (Varianz) definiert. Sie ist der wichtigste quantitative Wert für Variabilität der Daten und für weitere statistische Methoden (z. B. statistische Testverfahren) notwendig.

In dem Beispiel soll nun überprüft werden, ob Frauen und Männer die Stadtbusse unterschiedlich häufig nutzen.

			Geschlecht		Summen
			Männlich	Weiblich	
Wie oft nutzen Sie die Stadtbusse?	5–7 Tage pro Woche	Häufigkeit	176	296	472
		Prozent	37,3%	62,7%	100,0%
	3–4 Tage pro Woche	Häufigkeit	64	122	186
		Prozent	34,4%	65,6%	100,0%
	1–2 Tage pro Woche	Häufigkeit	33	53	86
		Prozent	38,4%	61,6%	100,0%
	1–3 Tage im Monat	Häufigkeit	12	28	40
		Prozent	30,0%	70,0%	100,0%
	Seltener	Häufigkeit	11	18	29
		Prozent	37,9%	62,1%	100,0%
Summen		Häufigkeit	296	517	813
		Prozent	36,4%	63,6%	100,0%

Tabelle 3-4: Beispiel Kontingenztabelle »Wie oft nutzen Sie die Stadtbusse?«

Die statistische Datenanalyse bietet die Möglichkeit, Zusammenhänge zu beschreiben, das geschieht durch die Berechnung des so genannte Korrelationsmaßes. Gibt es beispielsweise einen Zusammenhang zwischen dem Geschlecht, dem Alter, dem Autobesitz und der Nutzung des Stadtbusses? Das Maß des Zusammenhangs wird entsprechend des Skalenniveaus der untersuchten Variablen kategorisiert.

Die bisher vorgestellten Auswertungen beziehen sich nur auf die tatsächliche Stichprobe. Wie können aus diesen Daten Rückschlüsse auf ganze Populationen, z. B. auf die Einwohner eines Stadtteils oder der ganzen Stadt gezogen werden? Dafür verwendet man Inferenzstatistik (schließende Statistik). Die schließende Statistik stellte Hypothesen auf, die für die gesamte Population gelten sollen und prüft diese Gültigkeit für die Stichprobe. Voraussetzung ist, dass die Stichprobe durch eine Zufallsauswahl gewonnen wurde (siehe Kapitel 2.4 auf Seite 23). Nur bei Zufallsstichproben haben die statistischen Einheiten einer Population – wie Kugeln in einem undurchsichtigen Behälter – eine berechenbare Wahrscheinlichkeit, als Teil der Stichprobe gewählt zu werden. Dabei muss jedes Element der Population die gleiche Chance haben, in das Auswahlverfahren einbezogen zu werden.

Bei der schließenden Statistik kommen verschiedene statistische Testverfahren zum Einsatz. Die Wahl hängt wieder vom Skalenniveau ab. Manche Testverfahren erfordern beispielsweise eine spezielle Verteilung (z. B. Normal-/Gauß-Verteilung). Für mobilitätsbezogene Evaluationsverfahren ist oft die Korrelation zwischen zwei oder mehr Gruppen von Bedeutung. Die am häufigsten verwendeten Testverfahren in diesem Zusammenhang sind der t-Test und der Chi-Quadrat-Test, sie werden in den nächsten Schritten vorgestellt.

Nullhypothese und Alternativhypothese

Statistische Testverfahren untersuchen die Frage, ob die zufallsbedingte Streuung der erhobenen Messwerte ausreicht, um Effekte bzw. Zusammenhänge zu erklären. Es wird die Nullhypothese (H_0) aufgestellt, wenn in der Grundgesamtheit eine Gleichheit von Sachverhalten angenommen wird, dass heißt, es keine Effekte oder Zusammenhänge mit einer Maßnahme oder anderen Variablen gibt. So wird von der Nullhypothese ausgegangen, wenn z. B. eine Mobilitätsmaßnahme keine Wirkung zeigt oder wenn sich zwischen Bewohnern aus ausgewählten Wohnbezirken und deren Verkehrsmittelwahl

kein erkennbarer Zusammenhang ableiten lässt. Das Ergebnis der Stichprobe ist dann auf eine zufallsabhängige Streuung zurückzuführen. Die Alternativhypothese (H_1) besagt, dass der in der Stichprobe gefundene Zusammenhang bzw. Effekt in der Grundgesamtheit existiert. Die Nullhypothese ist die Grundlage der Teststatistik. Gibt es deutliche Widersprüche zwischen Nullhypothese und Messergebnissen, wird erstere zugunsten der Alternativhypothese verworfen.

Wird die Aussage in einer Hypothese mit einer Richtung formuliert, spricht man von einer gerichteten Hypothese. Eine ungerichtete Hypothese liegt dann vor, wenn nur eine Differenz angegeben wird. Diese Differenzierung enthält keine Bewertung der beiden Möglichkeiten (siehe Tab. 3-5 für Beispiele).

	Alternativhypothese H_1	**Nullhypothese H_0**
Gerichtete Hypothese (einseitiger Test)	Die wahrgenommene Sicherheit rund um Schulen mit einheitlich gestalteter Schulumgebung ist höher als bei Schulen mit unveränderter Umgebung.	Die wahrgenommene Sicherheit rund um Schulen mit einheitlich gestalteter Schulumgebung ist gleich oder niedriger als bei Schulen mit unveränderter Umgebung.
	Regelmäßige Busnutzende bewerten den Sitzkomfort im Stadtbus niedriger als Seltennutzende.	Regelmäßige Busnutzende und Seltennutzende bewerten den Sitzkomfort gleich.
Ungerichtete Hypothese (zweiseitiger Test)	Es gibt einen Zusammenhang (Korrelation) zwischen der wahrgenommenen Sicherheit und der Art der Schulumgebung.	Es gibt keine Korrelation zwischen der wahrgenommenen Sicherheit und der Art der Schulumgebung.
	Die Bewertungen des Sitzkomforts im Stadtbus durch regelmäßige Busnutzende unterscheidet sich deutlich von den Bewertungen der Wenignutzende.	Die Bewertungen des Sitzkomforts im Stadtbus durch regelmäßige Busnutzende unterscheidet sich nicht deutlich von den Bewertungen der Wenignutzende.

Tabelle 3-5: Beispiele für gerichtete und ungerichtete Hypothesen

Bestimmung des Signifikanzniveaus

Für die Annahme oder das Verwerfen der Nullhypothese ist die Bestimmung des Signifikanzniveaus wichtig. Eine Nullhypothese wird nur zurückgewiesen, wenn die Fehlerwahrscheinlichkeit kleiner oder gleich 5% oder 1% ist. Diese Grenze wird als Alpha-Fehler oder Signifikanzniveau bezeichnet. Führt eine Umfrage zu einer in der beschriebenen Weise begründeten Schlussfolgerung, wird von einem signifikanten ($\alpha = 5\%$) oder sehr signifikanten ($\alpha = 1\%$) Ergebnis gesprochen. Ein nicht signifikantes Ergebnis beweist jedoch nicht die Gültigkeit der Nullhypothese. Sie belegt nur, dass die Richtigkeit der Alternativhypothese nicht bewiesen werden konnte.

Wahl des richtigen Testverfahrens für die vorliegenden Daten

In diesem Schritt wird die geeignete Berechnungsmethode auf der Basis der zugrunde liegenden Fragestellung (und des Skalenniveaus ausgewählt). Der Chi-Quadrat-Test stellt die geringsten Anforderungen bezüglich des Messniveaus. Daher eignet er sich besonders für Daten auf einer Nominalskala. Ein typischer Fall für einen Chi-Quadrat-Test ist eine vergleichende Analyse zwischen Männern und Frauen in Bezug auf die Frage: Wie oft nutzen Sie die Stadtbusse?

Der t-Test ist eine Methode, um Hypothesen über die Ungleichheit von zwei Durchschnittswerten (z. B. Bewertung des Sitzkomforts vorher-nachher) zu validieren. Die Variablen müssen zur Verwendung dieses Tests intervallskaliert sein. Der t-Test kann nur verwendet werden, wenn Mittelwert und Varianz auf der Grundlage des Skalenniveaus sinnvoll interpretiert werden können.

Berechnung der Testparameter

Mit Hilfe der Daten wird die einseitige oder zweiseitige Überschreitungswahrscheinlichkeit (siehe auch gerichtete und ungerichtete Hypothese) berechnet. Je nach Testmethode wird dieser Parameter t, U, Chi-Quadrat usw. genannt.

Testentscheidung

Die ermittelte Überschreitungswahrscheinlichkeit wird mit dem definierten Signifikanzniveau verglichen. Haben wir ein signifikantes Ergebnis, d. h. ist der Testparameter kleiner 5%, wird die Nullhypothese zugunsten der Alternativhypothese abgelehnt.

Interpretation

Zum Schluss erfolgt die Dateninterpretation. Dabei ist es sehr wichtig, keine übereilten Schlüsse zu ziehen und über das Ziel hinaus zu schießen. Eine Dateninterpretation erfordert hohes Verantwortungsbewusstsein und Fingerspitzengefühl.

3.3 Qualitative Datenanalyse

Die Analyse und Interpretation von qualitativen Daten, die beispielsweise als aufgezeichnete Einzel- und Fokusgruppeninterviews, Fotos, Videoaufnahmen oder veröffentlichte Artikel vorliegen, erfordert ein anderes Know-how als der Umgang mit Zahlen und Häufigkeiten aus Umfragedaten. Daher erläutert dieses Kapitel die wichtigen Grundlagen für die Verarbeitung qualitativer Daten. Ziel ist es, in den Daten systematische Muster zu entdecken und diese zu erklären. Hilfreich für diese Auswertung können auch so genannte Metadaten, beispielsweise Notizen über den Kontext von Interviews (Datum, Ort, Besonderheiten der Befragungssituation, erste Ideen für die Analyse usw.) und relevante Informationen über die Interviewpartner (biografische Daten usw.) sein.

Transkription und Exzerpt der aufgezeichneten Daten

Bevor mit der Auswertung und Analyse der qualitativen Daten begonnen werden kann, müssen diese in eine einheitliche schriftliche Form gebracht werden. Diese Übertragung kann entweder detailliert Wort für Wort erfolgen (Transkription) oder durch Zusammenfassen der wesentlichen Inhalte in Stichworten (Exzerpt). Eine vollständige Transkription ist sehr zeitaufwändig, daher lohnt ggf. das Hinzuziehen von spezialisierten Dienstleistern. So kann die Transkription von einem einstündigen aufgezeichneten Interview leicht die vier- oder fünffache Zeit dauern.

Ein Transkript ist dann sinnvoll und nützlich, wenn die angesprochenen Themen sehr neu oder komplex sind und eine detaillierte Analyse notwendig ist. Es wird häufig bei biografischen Interviews verwendet. In der Regel werden während einer Evaluation keine übermäßig große Anzahl solcher Interviews durchgeführt.

Beim Exzerpt wird der Inhalt des Interviews bei der Datenerhebung zusammengefasst. Damit kann die Analyse jedoch nicht mehr scharf von der Interpretation getrennt werden. Dies ist auch der Fall, wenn bereits während des Interviews nur Stichworte notiert werden. Ein Exzerpt ist dann angemessen, wenn das Interview gut strukturiert war, die Antworten zu den besprochenen Themen präzise und klar waren und wenn überwiegend allgemeine Schlussfolgerungen gezogen werden sollen.

Bei beiden Methoden sollte darauf geachtet werden, die Daten hinreichend zu anonymisieren, dass also Namen oder andere personen- oder ortsbezogenen Informationen mit einem Code versehen werden. Wie schon beim Fragebogen sollte eine ID-Nummer vergeben werden. Bei Audiodateien lohnt

es sich auch, entsprechende Notizen zur Position auf der Aufnahme (z. B. Datum und Uhrzeit oder Bandabschnitt) zu machen, damit später z. B. ausgewählte Zitate leichter wieder auffindbar sind.

Codierung der Texte

Bei der Analyse qualitativer Daten sollen Muster erkannt werden. Damit dies gelingt, gibt es zunächst zwei Möglichkeiten, Auswertungskategorien dafür zu entwickeln. Im ersten Fall wird das Kategoriensystem anhand der Daten entwickelt, im zweiten Fall wird dieses schon vor der Datenerhebung festgelegt. Letzteres funktioniert immer dann gut, wenn die Thematik schon im Vorfeld hinreichend bekannt ist.

Ein Kategoriensystem muss möglichst eindeutig beschrieben werden. Jede Kategorie wird definiert und mit Beispielen unterlegt. Wenn nötig, werden Abgrenzungen zu anderen Kategorien gemacht.

Diese Codierung kann manuell über Tabellenkalkulationssoftware wie beispielsweise Excel oder Word-Tabellen oder mit professioneller Software für qualitative Daten, beispielsweise ATLAS.ti, NVIVO oder MAXQDA, erfolgen. Ziel ist es, die Organisation der Textdaten zu erleichtern und die weitere Datenanalyse und -interpretation vorzubereiten. Ein Beispiel für ein Kategoriensystem, einen sogenannten Codebaum, wird in der Abbildung 3-3 auf der linken Seite dargestellt. Auf der rechten Seite sieht man, dass einzelne Textstellen den Codes zugeordnet sind.

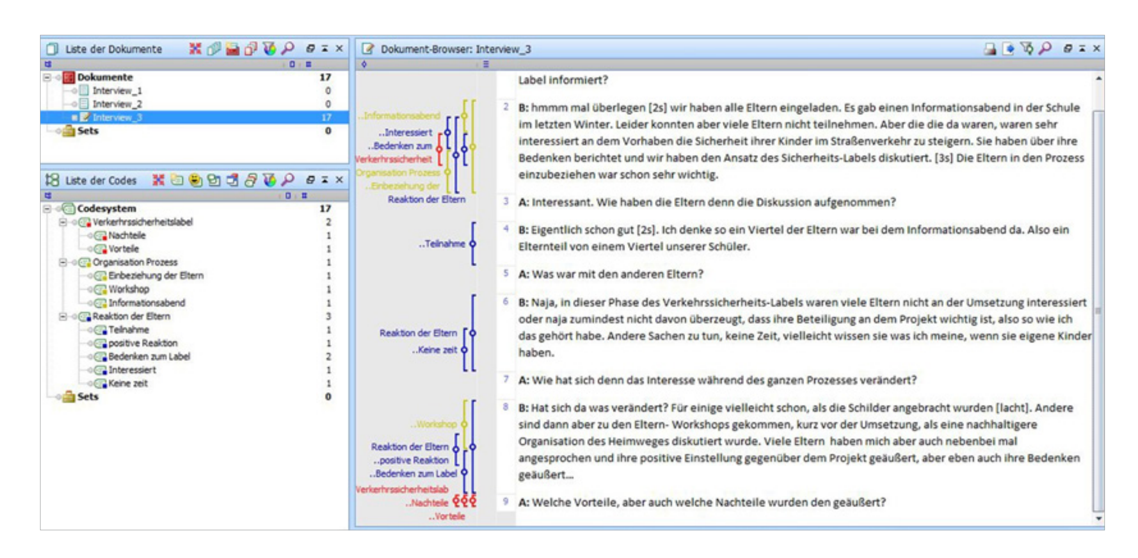


Abbildung 3-3: Visualisierung der Codierung einer transkribierten Befragung

Die Software bietet Unterstützung bei der Auswertung, aber nur für diejenigen, die auf strukturierte Art und Weise vorgehen. Die Auswertung sollte durch Notizen gut dokumentiert sein – jeder Schritt der Analyse muss auch später gut nachvollziehbarsein.

Um die Objektivität und Gültigkeit (Interrater-Reliabilität) zu gewährleisten, wird dringend empfohlen, dass mindestens zwei Personen voneinander unabhängig an den Texten arbeiten und die Codierung vornehmen. Weicht die Codierung zwischen den beteiligten Codierenden stark voneinander ab, sollten diese die unterschiedliche Bewertung diskutieren, ggf. das Kategoriensystem überarbeiten und/oder ihre Art der Codierung vereinheitlichen. Diesen Vorgang nennt man kommunikative-Validierung.

Kasten 11: Codieren der geführten Interviews

Eine praktische Möglichkeit kann sein, zuerst einen groben Entwurf der Auswertungskategorien zu erstellen, den sogenannten Codebaum. Dieser deckt alle Themen ab, die für die Evaluationsarbeit im Projekt wichtig erscheinen. Der Entwurf sollte sich an dem Leitfaden für die Befragungen orientieren. Nachdem die ersten transkribierten Befragungstexte codiert wurden, kann die Struktur gemäß den tatsächlichen Antworten neu organisiert werden. Dazu können auch Unterkategorien (Untercodes) angelegt werden. Es ist notwendig, den Codebaum flexibel zu handhaben, weil möglicherweise bei der Codierung der zweiten oder dritten Befragung eine wichtige neue (Unter-)Kategorie entstehen kann, für die ein neuer (Unter-)Code erforderlich ist. Mit diesen neuen Erkenntnissen, die während der Codierung gewonnen werden, sollten rückwirkend auch die bereits codierten Befragungen dahingehend geprüft werden, ob weitere relevante Informationen für diese Kategorie vorhanden sind. Die Codierung ist somit Bestandteil der Analyse des Materials und spiegelt den Denkprozess wider.

Tiefe und Genauigkeit der Codierung sind immer abhängig von Zweck und Tiefe der Interviews. Im Folgenden werden Anregungen für die Bildung von Kategorien gegeben:

- Ausgearbeitete oder angewendete Strategien, um ein Ziel zu erreichen
- Bedeutungen, Interpretation, und Relevanz von Phänomenen und Ereignissen (Einfluss der Handlungen der Befragten)
- Beziehungen oder Interaktionen zwischen Menschen oder Institutionen
- Ereignisse, die zeigen, was (und wie) etwas passierte
- Handlungen und Verhaltensweisen, die das repräsentieren, was gesagt oder getan wurde
- Handlungen, die mit einem bestimmten Hintergrund vorgenommen wurden (auch mit anderen beteiligten Personen)

Vergleich und Verallgemeinerung

Im nächsten Schritt werden die Antworten oder Fälle miteinander verglichen und wieder in Beziehung zu den Evaluationsfragen gesetzt.

Auf der Ebene des Einzelfalls (Einzelbefragung) kann geprüft werden, ob die Antworten eines Interviewten über die gesamte Befragung hinweg konsistent bleiben. Innerhalb einer Kategorie können Vergleiche z. B. zur Varianz des Themas angestellt werden. Und der Vergleich zwischen den Interviews zeigt z. B. unterschiedliche Handlungsalternativen oder subjektive Wahrnehmungen auf.

Auch hier lohnt es sich, bei der Arbeit laufend kleine Notizen zu Interpretationen zu machen und diese mit den Notizen aus der Dateneingabe zu vergleichen.

Für die Berichtlegung kann es hilfreich sein, eine Zusammenstellung der Antworten in einer Tabelle vorzunehmen. Für einzelne Kategorien können auch relevante Zitate genutzt werden. Dies erleichtert den Leserinnen und Lesern das Verständnis.

Insbesondere bei Expertenbefragungen oder längeren Einzelinterviews kann es sehr nützlich sein, die interviewte(n) Person(en) um eine abschließende Prüfung und Kommentierung zu bitten. Dies kann Fehlinterpretationen vermeiden. Auch hier gilt es wieder, die Datenschutzrichtlinien stets zu berücksichtigen.

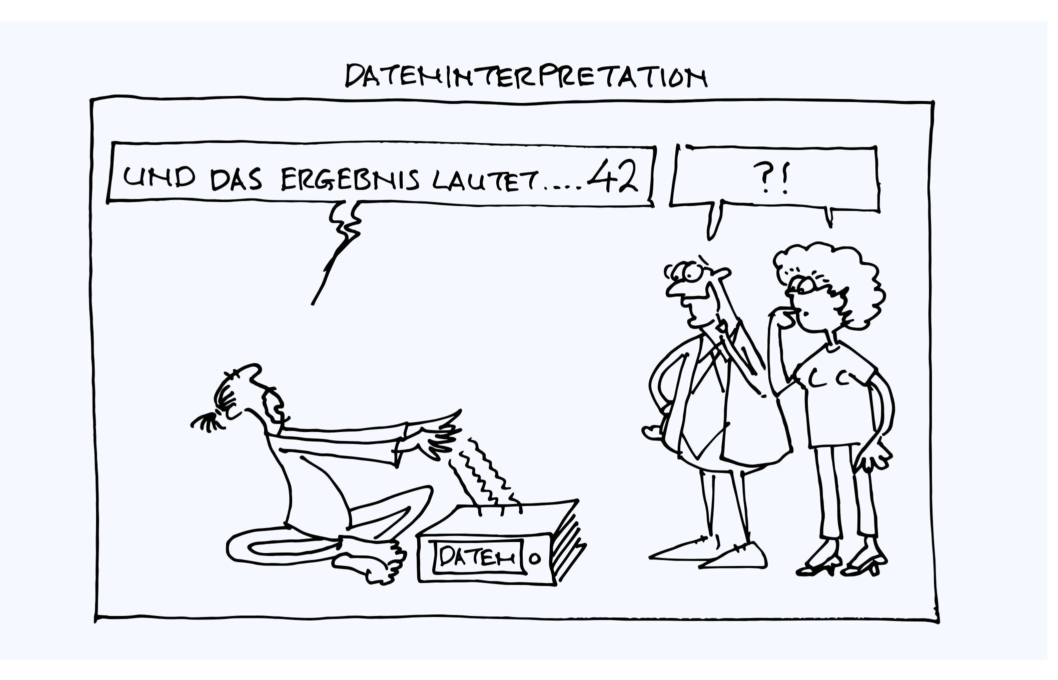
3.4 Ergebnisdiskussion

Was bedeuten alle diese Daten für die verkehrsbezogene Maßnahme, für die Stadt und ihre Bevölkerung? Der letzte Schritt der Auswertung gilt der Interpretation der Ergebnisse und dem Ausblick.

Es hat sich für Berichte (ob mündlich oder schriftlich) bewährt, die Präsentation der Daten von der Diskussion der Ergebnisse zu trennen. Zunächst werden also die relevanten Fakten möglichst objektiv durch Indikatoren, Messgrößen, Parameter und erklärende Texte dokumentiert (z. B. die Zahl der

Unfälle vor Schulen pro Jahr für die letzten 6 Jahre). Erst im nächsten Schritt (im Bericht ist das ein eigenes Kapitel), erfolgt eine Wertung der Ergebnisse, einschließlich der Einschätzung, in welchem Umfang die Ziele erreicht wurden und in welchem Maß die Maßnahmen zur Zielerreichung beitrugen. Der Grund für eine voneinander getrennte Datenpräsentation und Interpretation liegt darin, dass es den Leserinnen oder Zuhörern möglich sein soll, aus dem Datenmaterial eigene Schlüsse zu ziehen.

Eine gut ausgearbeitete und vorsichtig-kritische Interpretation der Ergebnisse vermeidet Missverständnisse bei internen und externen Interessengruppen und macht die Daten weniger anfällig für Prüfende und Kritisierende. Dazu müssen die Herangehensweise, die Argumente und Schlussfolgerungen nachvollziehbar, verständlich und logisch dargestellt werden.



An der Stelle ist der Bezug zu den Zielen der Maßnahme sehr wichtig (siehe auch Kapitel 2.1 auf Seite 15). Führte die Maßnahme dazu, dass die gewünschten und durch Indikatoren definierten Effekte eingetreten sind? Unterscheiden sich die Effekte von denen, die in dem Referenzszenario eingetreten sind oder wären? Wo oder für wen war die Maßnahme erfolgreich, wo nicht? Können die Ergebnisse auch durch andere Daten oder Beobachtungen gestützt werden (Triangulation, siehe Kasten 12)? Gab es eventuell weitere Einflussfaktoren, die das Ergebnis begünstigten, welche aber bei der Planung der Evaluation nicht berücksichtigt wurden?

Falls sich die gewünschten Ergebnisse oder Effekte der Maßnahme nicht zeigen, ist eine ebenso sorgfältige und ehrliche Auswertung und Diskussion nötig. Welche Erklärungen oder Gründe gibt es?

Kasten 12: Triangulation

Die in der Evaluation verwendeten Daten stammen normalerweise aus mehreren Quellen und die Ergebnisse der Datenanalyse werden gegeneinander abgewogen oder miteinander verglichen. Dieses Verfahren wird Triangulation genannt, kann die Zuverlässigkeit der gewonnenen Daten sichern und methodisch richtige Schlussfolgerungen unterstützen. Die Triangulation wird für Datenerhebungsmethoden (Unterschiede bei den angewendeten Methoden), aber auch für Informationsquellen (Sammeln von Informationen aus unterschiedlichen Befragungsgruppen) verwendet. Durch die Kombination mehrerer Methoden und Datenarten können die Nachteile und Unschärfen bei den einzelnen Methoden ausgeglichen werden.

4 Prozessevaluation

Im Gegensatz zur Wirkungsevaluation, die erst nach Umsetzung der Maßnahme komplettiert werden kann, begleitet die Prozessevaluation die gesamte Maßnahme, einschließlich der Entwicklungsphase, von Anfang an. Das Ziel ist somit, bereits in einem sehr frühen Stadium der Planung schrittweise alle Abläufe zu durchleuchten, um Hindernisse und Schwierigkeiten während der Umsetzungsphase und darüber hinaus möglichst zu minimieren. Im Idealfall ist die Prozessevaluation also keine zusätzliche Aktivität, sondern ein fester Bestandteil einer Maßnahme.

Besonders für die Maßnahmen der kommunalen Verkehrsplanung gilt, dass häufig sehr viel Zeit zwischen der ersten Idee und der tatsächlichen Umsetzung oder Inbetriebnahme liegt. Oft vergehen Jahre. Es ist also normal, dass sich im Laufe des gesamten Verfahrens beispielsweise die politischen Rahmenbedingungen und damit möglicherweise auch strategische Ziele ändern können. Auch andere Bereiche sind oft von Änderungen betroffen: Veränderungen im kommunalen Haushalt, oft einhergehend mit einer Verknappung der finanziellen Mittel, Veränderungen beim Personal und den beteiligten Fachleuten oder soziale oder gesellschaftliche Veränderungen treten immer wieder auf. Hinzu kommen auch technische Innovationen, die einen großen Einfluss auf die Mobilität haben können. Allein die Einflüsse der mobilen Applikationen (»Apps« und mobiles Internet), die heute auf Smartphones genutzt werden, haben das Mobilitätsverhalten einer breiten Bevölkerung innerhalb der letzten fünf Jahre sehr stark beeinflusst. Ob freie Parkplätze, die schnellste U-Bahnverbindung, der sicherste Radweg, attraktive Fußgängerzonen, Geschäfte, Restaurants und Cafés, Mitfahrgelegenheiten, Taxis, Mietwagen, Car Sharing oder Call a Bike – dies alles und noch vieles mehr lässt sich über Apps online schnell finden oder nutzen. Zudem bieten diese meistens auch die Möglichkeit, direkt Reservierungen oder Buchungen vorzunehmen und gleich zu bezahlen. Haben wir diese gravierenden Veränderungen und Erleichterungen durch die technischen Neuerungen im Jahr 2000 vorhergesehen?

4.1 Ziele und Ergebnisse der Prozessevaluation

Die Umsetzung von mobilitätsbezogenen Maßnahmen ist eine zeitaufwändige, oft teure und manchmal mühsame Aufgabe, die mit vielen Akteuren abgestimmt und durchgeführt werden muss. Sie sollten daher eine möglichst große Wirkung erzielen. Viele unterschiedliche Faktoren werden sich im Laufe der Planung und Umsetzung als fördernd oder auch als hemmend herausstellen. Es gilt, diese bereits im Vorfeld zu erkennen und die Folgen bei der Planung in angemessener Weise zu berücksichtigen. Die Prozessevaluation hat daher zum Ziel, Erfolge der Maßnahmen so weit wie möglich zu sichern. Sie dient also nicht allein der Bewertung von Prozessen, sondern auch dazu, möglichst effiziente Vorgehen und Verfahren im Prozess selbst zu finden.

Bei der Prozessevaluation werden daher – fortlaufend über den gesamten Prozess – die folgenden Aspekte untersucht:

- Ziele und Zielfindungsprozesse auf allen Ebenen
- Indikatoren und ihre Auswahl
- Entwicklung des Wirkungsmodells, u. a. Ursache-Wirkungsketten
- Planung und Entwicklung der Maßnahme
- Umsetzung der Maßnahme, einschließlich Bau- oder Herstellungsphase, Probebetrieb und Betrieb

Das Ergebnis der Prozessevaluation ist eine Dokumentation der während des Projekts eingetretenen Ereignisse. Sie hilft, Entscheidungen und Lernerfahrungen nachvollziehbar zu machen – für Prozessbeteiligte und für weitere Interessierte. Dies kommt dem Projektablauf und der Zielerreichung zugute. Sie stellt außerdem fundierte Argumente für die politischen Entscheidungen, die Verwaltungen und die Öffentlichkeit bereit. Nicht zuletzt liefert die Prozessevaluation auch wichtige Ergebnisse über den Erfolg einer Maßnahme. Wenn es beispielsweise darum geht, ob das Projekt in einer anderen Stadt umgesetzt werden soll, ist es schließlich nicht nur wichtig zu wissen, welche Auswirkungen vom

Projekt erwartet werden können, sondern auch, wie der Prozess abläuft, um das Ziel zu erreichen und welche Bedingungen für eine erfolgreiche Durchführung gegeben sein müssen.



4.2 Widerstände und Treiber

Die Prozessevaluation sucht nach Einflüssen, die das Projekt oder die Maßnahme entweder fördern oder behindern. Im Folgenden soll dabei insbesondere auf die an der Maßnahme beteiligten Personen eingegangen werden. Deren Unterstützung, Teilnahme, Arbeitsweisen, Erfahrungen, Wahrnehmungen und Einschätzungen, aber auch deren Kritik, Sorgen, Ängste, Unwillen, Ablehnung oder Tatenlosigkeit haben einen großen Einfluss auf ein Projekt und dessen Verlauf. Es kann daher für ein Projekt oder eine Maßnahme sehr wichtig sein, bereits im Vorfeld Haltungen unterschiedlicher Personengruppen zu analysieren und in Bezug auf deren Wirkung auf das Projekt abzuschätzen.

Als Widerstände oder Hürden für den Prozess können sich alle Ereignisse oder Konstellationen erweisen, die sich dem Prozess beim Erreichen der Ziele der Maßnahme in den Weg stellen. Neben wirtschaftlichen oder sozialen Faktoren können es auch etwa organisatorische, strukturelle, politische oder kulturelle Aspekte sein. Treiber sind dagegen die Ereignisse oder Bedingungen, die den Prozess beim Erreichen der Ziele der Maßnahme unterstützen.

Für eine gründliche Prozessevaluation kann es hilfreich sein, alle vermuteten, fördernden und hemmenden Einflüsse in einer Tabelle aufzulisten und diese gegeneinander abzuwägen (siehe Tabelle 4-1). Es kann dabei durchaus vorkommen, dass der gleiche Einflussfaktor sowohl eine fördernde als auch eine hemmende Wirkung zeigt.

Bereich	Beispiele für Widerstände	Beispiele für Treiber
Politisch/ strategisch	Konflikte zwischen (politischen) Interessengruppen aufgrund unterschiedlicher Einschätzungen bezüglich Problemen und deren Lösungen. Opposition wichtiger Akteure aufgrund politischer und/ oder strategischer Motive, fehlende Leitbilder für eine nachhaltige Mobilität	Engagement wichtiger Akteure aufgrund politischer und/oder strategischer Motive, breite Akzeptanz der Leitbilder für eine nachhaltige Mobilität, hohe Kompromissbereitschaft der beteiligten Akteure
Institutionell	Komplexe Verwaltungsstrukturen, Verfahren und Routinen, behindernde Gesetze, Regeln, Vorschriften und deren Anwendung, hierarchische Strukturen von Organisationen und Programmen, viele unterschiedliche Ansprechpartner und Verantwortliche	Effiziente Verwaltungsstrukturen, Verfahren und Routinen, erleichternde Gesetze, Regeln, Vorschriften, klare Verantwortlichkeiten, Kompetenz und Erfahrung bei den Verantwortlichen
Kulturell	Kulturelle Gegebenheiten und Lebensstile, mangelndes Verständnis oder Toleranz, wenig Erfahrung mit partizipativen Prozessen	Kulturelle Gegebenheiten und Lebensstile, Vorhandensein von Trendsettern und Vorbildern, aktive Bürgerschaften, Eigeninitiative, Sprechen einer gemeinsamen Sprache, Offenheit und Toleranz
Engagement, Kommunikation	Unzureichende Beteiligung oder Sensibilisierung von wichtigen (politischen) Interessengruppen, unzureichende Beratung, Beteiligung oder mangelndes Bewusstsein von Bürgern oder Nutzerinnen	Konstruktive und offene Beteiligung von wichtigen (politischen) Interessengruppen, konstruktive und offene Mitsprache, Engagement, lebhaftes Beteiligungsverhalten, Pressefreiheit, Vorhandensein verschiedener Kommunikationskanäle
Planung	Unzureichende technische Planung und Analyse zur Feststellung der Anforderungen bei der Umsetzung der Maßnahme, unzureichende Ressourcenplanung und Marktanalyse, begrenztes Verständnis der Anforderungen der Nutzerinnen und Nutzer	Kompetenz und Erfahrung bei den Planern, präzise technische Planung und Analyse zur Feststellung der Anforderungen, solide Wirtschaftsplanung und Marktanalyse, sorgfältige Bedarfsanalyse und gutes Verständnis der Anforderungen der Nutzerinnen und Nutzer
Organisatorisch	Unklare Organisationsstrukturen und Verantwortlichkeiten, mangelndes Zeitmanagement, fehlende oder unzureichende Partnerschaftsvereinbarungen, mangelnde Führung, fehlende individuelle Motivation oder fehlendes Know-how von wichtigen Beteiligten der Maßnahme	Konstruktive Partnerschaftsvereinbarungen, starke und klare Führung, hoch motivierte Beteiligte der Maßnahme als »Lokalmatadoren«, enge und vertrauensvolle Zusammenarbeit zwischen den beteiligten Verwaltungsabteilungen und Partnern
Finanziell	Zu große Abhängigkeit von öffentlichen Geldern und Subventionen, mangelnde Bereitschaft der Wirtschaft, sich finanziell zu beteiligen	Verfügbarkeit von öffentlichen Geldern und Subventionen, Bereitschaft der Wirtschaft sich finanziell zu beteiligen
Technisch	Zusätzliche technische Anforderungen, Technologie noch nicht verfügbar, technische Probleme	Neues Potenzial durch Technologien, neue Technologien verfügbar, Nutzerfreundlichkeit

Tabelle 4-1: Beispiele für Widerstände und Treiber

Die in Tabelle 4-1 genannten Kategorien sind natürlich nicht vollständig, es gibt noch viele weitere. Für jedes Projekt sollten deshalb nicht nur diese vorgegebenen Bereiche untersucht werden, sondern in einem weiteren Schritt möglichst knapp aber präzise und auch für Außenstehende nachvollziehbar weitere Details die eine Auswirkung auf den Prozess haben, beschrieben werden. Fragen

sind dabei: Wie ist die Situation genau? Wie sind die Abläufe und Zusammenarbeiten geregelt? Was oder wer fördert oder könnte behindern? Welche Möglichkeiten sehen die Beteiligten, die Abläufe zu optimieren?



Es ist möglich, dass ein Projekt aufgrund verschiedener Rahmenbedingungen schneller vorankommt, als geplant. Abläufe vollziehen sich reibungslos, die Pilotphase kommt bei der Bevölkerung gut an, die technische Umsetzung ist leichter als gedacht oder das Wetter begünstigt die Ausführung. Viele dieser Treiber können, nachdem sie erkannt und erfasst wurden, noch stärker für die Erreichung der Projektziele nutzbar gemacht werden.

Jedoch tauchen in der Praxis deutlich mehr Hindernisse auf, als im Vorfeld bedacht werden konnten. Das Projekt verzögert sich, kommt nur schleppend voran und alles dauert länger als geplant. Wichtige Praxispartner können aussteigen und manchmal wird die Finanzierung eng. Die Öffentlichkeit reagiert zögerlich, die Presse kritisch. Folglich geraten die Verantwortlichen unter Druck und eine erfolgreiche Umsetzung des Projekts wird gefährdet.

Zudem ist ein vorausschauendes Erkennen und Bewerten von Widerständen und Treibern eine wichtige Abschätzung von Risiken. Eine begleitende Prozessevaluation kann dann auf fast alle sich verändernden Rahmenbedingungen zeitnah reagieren. Fortschritte, aber auch Verzögerungen, werden rechtzeitig erkannt, so dass die Projektplanung entsprechend angepasst werden kann. Auch hier hilft eine nachvollziehbare Dokumentation, die Kommunikation nach außen entscheidend zu verbessern. Dies sind die stärksten Argumente für eine Prozessevaluation.

Auch nachdem die ersten Hürden und Hemmnisse überwunden wurden und das Projekt wieder wie geplant läuft und nachdem positive Ereignisse ausgewertet wurden, sollte die Prozessevaluation weitergeführt werden. Einerseits ist es motivierend für alle Beteiligten, von den Erfolgen zu hören und für ihre Arbeit gelobt zu werden, andererseits können jederzeit wieder neue Hindernisse oder Treiber auftreten, die bewältigt werden müssen oder die nutzbar gemacht werden können.

4.3 Methoden der Datenerhebung

Die Methoden für eine Prozessevaluation müssen vielfältigen Anforderungen gewachsen sein. Sie sind entwickelnd, beschreibend, kontinuierlich, flexibel und induktiv. Sie sollten möglichst in allen Phasen der Prozessevaluation einsetzbar sein. Zur Planung der entsprechenden Methoden sollte schon begleitend zur Maßnahmenplanung ein Zeitplan für die Prozessevaluation erstellt werden.

Nicht nur die Häufigkeit, auch die Intensität, Art und Umfang der Datenerfassung hängen in hohem Maße von der Verfügbarkeit von Personal und finanziellen Ressourcen ab. Die Evaluationsplanung sollte diese Wirtschaftlichkeits- und Effizienzaspekte genau wie bei einer Projektplanung berücksichtigen und die verfügbaren Methoden wohlüberlegt einsetzen. Dennoch gilt es, nicht am falschen Ende zu sparen. Es sollte in jedem Fall das Ziel sein, den gesamten Ablauf zu evaluieren und alle relevanten Aktivitäten in Bezug auf den Maßnahmenprozess zu dokumentieren – von der Vorbereitung über die Umsetzung bis zum Betrieb.

Es gibt viele erprobte Methoden für die Prozessevaluation. Im Folgenden werden Instrumente vorgestellt, die im besonderen Maße dazu geeignet sind, die verschiedenen Blickwinkel aller Beteiligten zu erfassen und für den Prozess nutzbar zu machen. Zudem können auch die Methoden aus der Wirkungsevaluation (siehe Kapitel 2.5 auf Seite 29) zum Einsatz kommen.

Standardisierte Formulare

Die Datenerhebung und Dokumentation wird enorm erleichtert, wenn standardisierte Formulare zum Einsatz kommen. Diese können für verschiedene Bereiche entwickelt werden. Vorteile sind, dass die jeweiligen Verantwortlichen oder Beteiligten diese selbst ausfüllen können und die Ergebnisse untereinander oder auch mit anderen Projekten verglichen werden können. Ein solches Formular sollte folgende Informationen abfragen:

Allgemeine Informationen	· Allgemeine (administrative) Informationen wie Ziele, Zielgruppen, Interessenvertreter, Partner und Maßnahmen
Widerstände	· Kurze und klare Beschreibung der wichtigsten Widerstände · Was ist geschehen? Wie kam es dazu? Welche Auswirkungen hat dies auf den Prozess der Maßnahmen?
Treiber	· Kurze und klare Beschreibung der wichtigsten Treiber · Was ist geschehen? Wie kam es dazu? Welche Auswirkungen hat dies auf den Prozess der Maßnahmen?
Aktionen	· Kurze und klare Beschreibung der wichtigsten Aktionen, die ergriffen wurden, die Treiber zu nutzen oder die Prozesswiderstände zu überwinden.
Risiko	· Abschätzung des Risikos zur Erreichung der Ziele
Sonstiges	· Kurze und klare Beschreibung der anderen relevanten Informationen, um den Maßnahmenprozess zu erklären: außergewöhnliche Bedingungen, Methoden, etc.

Abbildung 4-1: Schema eines standardisierten Formulars

Das Kernstück des Formulars sollte jedoch die Dokumentation der Widerstände und Treiber des Prozesses sowie die daraus abgeleiteten Aktivitäten sein. Es ist sicher auch hilfreich, dem Formular ergänzende Tabellen mit Kategorien für Treiber und Widerstände (wie in Kapitel 4.2 vorgestellt) beizulegen. Gibt es viele Widerstände oder Treiber, können diese auch noch subjektiv oder objektiv bewertet und/oder in eine Rangreihenfolge gebracht werden. Handelt es sich um ein längeres oder umfangreiches Projekt, sollten auf diese Weise alle möglichen Teilschritte oder Maßnahmen dokumentiert werden. Idealerweise findet die Bewertung nicht nur rückblickend, sondern zu mehreren Zeitpunkten des Projekts auch vorausschauend statt. Die Beurteilung eines Projektes durch standardisierte Formulare ist eine einfache und effektive Methode. Außerdem ergibt sich der Nebeneffekt, dass diese Zusammenstellungen einfach in den Abschlussbericht übernommen werden können. Allerdings bleibt bei dieser Methode der Umfang und die Tiefe der Informationen beschränkt. Die Nachvollziehbarkeit,

die Übertragbarkeit auf andere Projekte sowie der Erfahrungsaustausch sind also geringer als bei anderen Verfahren.

Storytelling und Learning History

Die Methode Storytelling bzw. Learning History basiert auf der Erkenntnis, dass durch das mündliche Erzählen besonders gut Erfahrungen geteilt und weitergegeben werden können. »Learning history« meint in diesem Zusammenhang also »aus der Geschichte« lernen. Beteiligen sich unterschiedliche Akteure an der Methode, kann die gleiche Projektgeschichte aus unterschiedlichen Blickwinkeln erzählt und damit nacherlebt werden. Das regt die Reflektion enorm an und die Erkenntnisse bleiben länger im Gedächtnis als bei den meisten anderen Methoden. Storytelling wurde als Methode in den späten 90ern von Forschenden an einem amerikanischen Institut (MIT) entwickelt.

Am besten lässt sich Storytelling im Rahmen von Projekttreffen oder einem Workshop anwenden. Alternativ kann die Methode aber auch in Form von Einzelinterviews durchgeführt, ausgewertet und zu einem Gesamtbericht zusammengefügt werden. Jeder oder jede Beteiligte erzählt nach dem gleichen Schema seine oder ihre Version der Projektereignisse. Das kann der chronologische Ablauf sein sowie wichtige strategische Entscheidungen, Fakten, Daten, Ursachen und Wirkungen, Rückschläge und Erfolge, Lernerfahrungen, persönliche – auch emotionale – Bewertungen mit einbeziehen. Die Erzählung kann begleitend visualisiert werden, z. B. mit Fotos, Bildern oder mit Hilfe der Metaplantechnik. So entsteht ein mehrdimensionales Bild der Gesamtsituation.



Storytelling kann im Projektverlauf mehrfach eingesetzt werden. Die Methode eignet sich insbesondere zur Reflektion kritischer Situationen, beispielsweise wenn es zu größeren Verzögerungen kam und ein vorzeitiges Maßnahmenende drohte. Durch den hohen Lerneffekt bei den Beteiligten bietet sie die Chance für das gemeinsame Entwickeln einer Lösung. Wie ein solcher Workshop ablaufen könnte, wird in der folgenden Darstellung gezeigt:

vorher	**Vorbereitung**	· Mindestens drei Wochen vorher: Identifizierung der Teilnehmenden und einer Moderatorin/eines Moderators, Einladung versenden, Zeitplan erstellen, Raum und Verpflegung buchen, Material zusammenstellen (ca. 2 Arbeitstage)
Workshop	**Schritt 1** Vorstellung	· Begrüßung, Vorstellungsrunde, Vorgehen und Ziele des Workshops erläutern, Regeln festlegen, wichtige Begriffe klären (Widerstände, Treiber und Aktionen) (ca. 20 min)
	Schritt 2 Ablauf erstellen	· Präsentation eines groben Ablaufplans, Diskussion und Ergänzung durch die Teilnehmer, Diskussion der Chronologie des Projektes, Erstellen der Zeitachse (ca. 30-60 min)
	Schritt 3 Widerstände	· Erstellen einer Liste der aufgetretenen Widerstände und Hemmnisse, Diskussion und Kategorisierung (ca. 45 min)
	Schritt 4 Treiber	· Erstellen einer Liste der Treiber, Diskussion und Kategorisierung (ca. 45 min)
	Schritt 5 Aktionen	· Erstellen einer Liste der Aktionen, die als Reaktion auf ein eingetretenes hemmendes, oder begünstigendes Ereignis folgten (ca. 45 min)
	Schritt 6 Diskussionen	· Reflektierende Diskussion zu den Lernprozessen (ca. 45 min)
nachher	**Dokumentation**	· Zusammenstellen der Ergebnisse in einem möglichst standardisierten Bericht, im Idealfall direkt nach dem Workshop (ca. 3-4 Stunden)
	Feedback, Verifikation, Verbreitung	· Versenden des Berichtentwurfs an alle Teilnehmenden mit der Bitte um Rückmeldung und ggf. Ergänzung (spätestens zwei Wochen nach dem Workshop) · Überarbeitung und Versenden des Schlussberichts (ca. 2-3 Stunden)

Abbildung 4-2: Ablaufplan eines Learning-History-Workshops

Die Durchführung eines Workshops sollte gründlich in allen Schritten geplant und durchdacht werden. Wer soll teilnehmen? Wer moderiert? Welche Aspekte sollen besprochen werden? Was könnte das Ergebnis sein? Wo drohen Konflikte? Welche ergänzenden Informationen oder Unterlagen müssen bereit liegen?

Damit sich die Teilnehmenden angemessen einbringen können, sollten an einem solchen Workshop nicht mehr als zehn bis zwölf Personen teilnehmen. Idealerweise kommen diese Akteure aus ganz unterschiedlichen Bereichen, damit sie viele Sichtweisen einbringen können. Als Grundlage für die Zusammenstellung dient ggf. die Liste der Zielgruppen und Projektbeteiligten.

Der Erfolg des Workshops hängt sehr von den Fähigkeiten des Moderators oder der Moderatorin ab. Es ist eine anspruchsvolle Aufgabe, die Diskussionen durch Fragen zu lenken, ohne diese inhaltlich zu beeinflussen, den Zeitplan einzuhalten, Kategorien zu bilden, und Ergebnisse festzuhalten und ggf. zu gewichten.

Kasten 13: Kompetenzen der Moderation für einen Learning-History-Workshop

Er/Sie sollte:

- Erfahren in Moderationstechniken und Gruppendiskussionen sein
- Fähig sein, eine vertrauensvolle Atmosphäre aufzubauen, in der jeder Teilnehmende respektiert wird und offen sprechen kann
- So gut über das Projekt informiert sein, dass ihm/ihr die wichtigsten Aspekte bekannt sind, z. B. die Maßnahmenziele, beteiligte Interessengruppen sowie mögliche kontroverse Punkte und Konflikte.
- Gut über die Rollen und Verantwortlichkeiten der Beteiligten innerhalb der Maßnahme informiert sein

Seine/ihre Aufgaben sind:

- Zeit und Ablauf einhalten
- Zwischenergebnisse zusammenstellen, Clustern, Ergebnisse festhalten
- Beiträge von allen Teilnehmenden gleichermaßen ermöglichen
- Eine offene Diskussionskultur durch Fragen anregen

Das gemeinsame Aufstellen und Visualisieren einer Chronologie des Projektes ist der erste wichtige Schritt im Workshop. Dazu bietet sich die Diskussion an Hand einer Zeitachse an. Alle wichtigen Projektphasen werden entlang der Achse eingetragen. Damit das Aufstellen nicht zu lange dauert, sollten die Zeitachse und der grobe Projektverlauf schon vor dem Workshop vorbereitet werden. Im Laufe des Workshops werden dann alle relevanten Ereignisse ergänzt. Das Ergebnis ist eine kommentierte Zeitachse, die für sich schon ein interessantes Analyseinstrument darstellen kann. Haben einzelne Beteiligte von Ereignissen nicht oder erst später erfahren? Haben sie diese eventuell unterschiedlich gewichtet, z. B. ob sie Chancen oder Risiken bergen? Wie schnell konnte reagiert werden? Gab es Zeiträume, in denen viel und andere, in denen nichts oder wenig passiert ist? Welchen Einfluss hatte das auf die Motivation der Beteiligten und auf den Erfolg der Maßnahme?

Das Abfragen von Widerständen und Treibern kann entweder in einer offenen Diskussion erfolgen oder aber auch durch die sogenannte Kartenabfrage. Die Teilnehmenden werden dabei gebeten, auf Haftnotizzetteln oder Moderationskarten knapp ihre Beiträge aufzuschreiben. Diese können an Moderationswänden gesammelt, gruppiert, gewichtet und diskutiert werden. Der Prozess des durch die Moderatoren angeleiteten Sammelns, der Strukturierung und der Diskussion wird bei den Teilnehmenden dazu führen, dass sie die einzelnen Aspekte aus unterschiedlichen Blickwinkeln und als Gesamtbild wahrnehmen können.

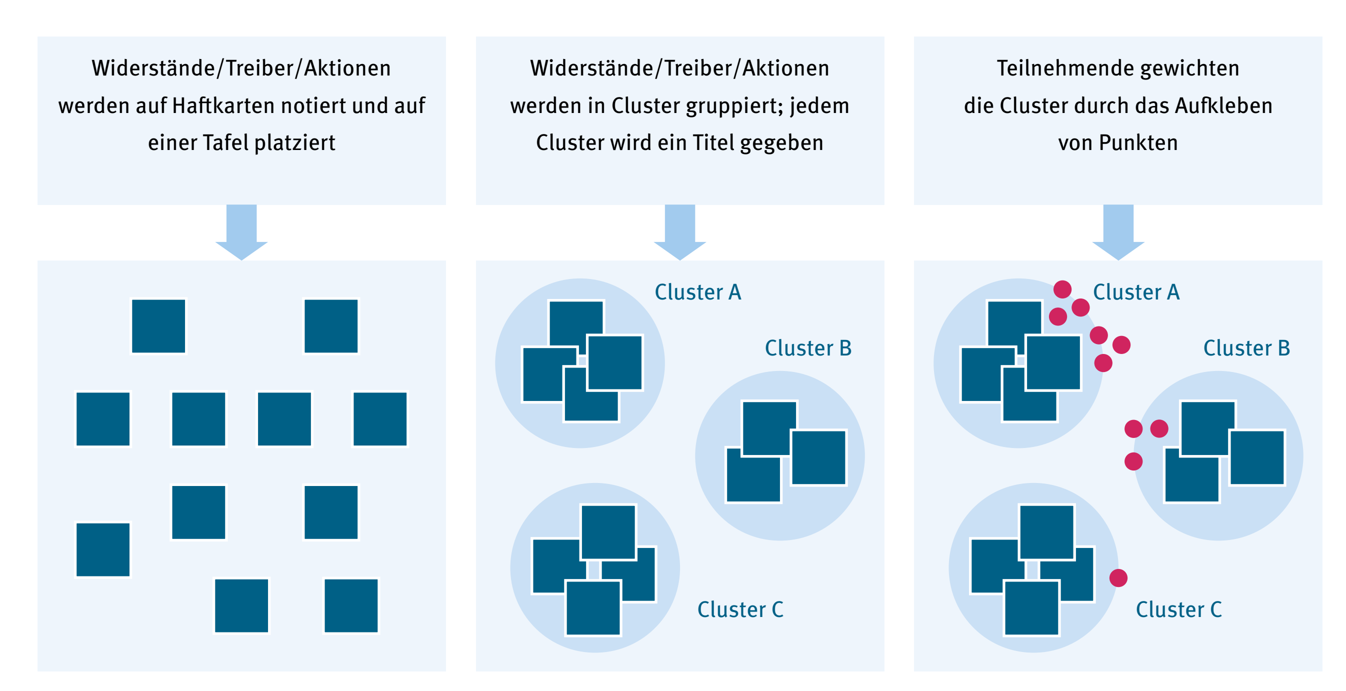


Abbildung 4-3: Beispiel für eine Kartenabfrage

Die umgesetzten Aktionen im Projekt können ebenfalls auf die Weise gesammelt und diskutiert werden. Eventuell ergibt es aber mehr Sinn, diese mit Hilfe anderer Moderationstechniken zu beschreiben und zu bewerten oder auf der Zeitachse zu dokumentieren.
Am Abschluss steht die zusammenfassende Diskussion. Dabei können alle Teilnehmenden für sich formulieren, welche Lernerfahrungen er oder sie im Projekt und im Workshop gemacht hat.
Zentrale Fragen zur Strukturierung der Diskussion können sein:

- Welche der Aktionen kann als Erfolg und welche als Misserfolg angesehen werden und warum?
- Was kann aus dem Projekt und seinem Verlauf gelernt werden? Was waren Erfolgsfaktoren? Was hat dem Projektverlauf geschadet? Wie können zukünftig Projekte und vor allem Prozesse dieser Art optimiert werden?

Auf der Grundlage der Ergebnisse wird ein Bericht erstellt und zunächst intern von allen Beteiligten geprüft, ergänzt und nach Überarbeitung wird er ggf. veröffentlicht.

Der große Vorteil der Methode des Storytelling-Workshops besteht darin, dass alle Beteiligten in einem Raum sind und sich zeitlich begrenzt konzentriert mit den gleichen Fragen beschäftigen. Das gegenseitige Zuhören, die Diskussion und das gemeinsame Bewerten unterstützen das Verständnis und fördern die Perspektivenübernahme. Die Methode kann sowohl in der Prozessevaluation als auch in der qualitativen Wirkungsevaluation eingesetzt werden. Es könnte aber sein, dass einige der Teilnehmenden auf die Einladung zu einem solchen Workshop zunächst zurückhaltend oder skeptisch reagieren. Sie empfinden die Methode als zu spielerisch oder zu »soft«. Durch eine neutrale Benennung kann solchen Vorurteilen aber gut vorgebeugt werden. Falls sich die Teilnehmenden gar nicht zum Mitmachen motivieren lassen, können die einzelnen Perspektiven auch getrennt voneinander in Einzelinterviews oder auch schriftlich eingeholt werden. Dabei kommen aber die beiden wichtigsten Elemente, nämlich das Zuhören und der Diskurs, deutlich zu kurz.

Sonstige Methoden

Es gibt noch weitere geeignete Methoden für die Prozessevaluation, beispielsweise Fokusgruppen oder Befragungen (siehe Kapitel 2.5 auf Seite 29). Beide werden auch in der Wirkungsevaluation eingesetzt. Wichtig ist, dass die Methoden geeignet sein sollten, das Verfahren selbst, den Verlauf, die Arbeitsabläufe – eben den Gesamtprozess mit allen fördernden und hemmenden Faktoren – zu bewerten. Es geht an dieser Stelle nicht um die Ergebnisse oder Effekte der Maßnahmen.

	Standardisierte Formulare	Learning History/ Storytelling	Fokusgruppe	Befragung
Wie?	Datenerhebung durch Ausfüllen eines standardisierten Formulars	Eingehende systematische Informationsbeschaffung und –bewertung mit allen Prozessbeteiligten	Moderierte Gruppendiskussion mit bis zu 15 ausgewählten Teilnehmenden bestimmter Zielgruppen zu einem speziellen Thema/Fokus	Strukturierte mündliche oder schriftliche Befragung
Einfache und effiziente Datenerhebung	+	+/-	-	+
Ausreichende Beteiligung der Interessengruppen	-	+	+/-	+/-
Liefert ein umfassendes Bild des Prozesses	+/-	benötigt erfahrene Moderation	benötigt erfahrene Moderation	benötigt erfahrene Interviewende
Möglichkeit, aus den erzählten »Geschichten« zu lernen	+	++	+/-	+/-
Angemessener Ressourcenaufwand	+	+/-	-	-
				Legende: ++ sehr gut; + gut; +/- ausreichend; - ungenügend

Tabelle 4-2: Methodenvergleich für die Datenerhebung in der Prozessevaluation

Natürlich hängt die Qualität der Ergebnisse in hohem Maß davon ab, wie gründlich die Methoden vorbereitet und durchgeführt werden.
Zusammenfassend gelten die folgenden Kriterien:

1. Die Methode sollte die unterschiedlichen Prozessbeteiligten erreichen, damit das Ergebnis ein möglichst objektives Bild ergibt.
2. Die Moderatoren müssen neutral sein, sie dürfen die Prozessbeteiligten nicht absichtlich oder unabsichtlich durch die Art der Fragen oder durch die Wahl der Methoden beeinflussen.
3. Der Aufwand für Vorbereitung, Durchführung und Berichterstattung sollte angemessen sein. Die Methode sollte bei den Beteiligten auf Akzeptanz stoßen. Ansonsten steigt die Gefahr des Scheiterns.
4. Bei einer Prozessevaluation kann es durch das Monitoring auch zu einer Kontrolle der am Projekt beteiligten Akteure kommen. Es sollte aber jederzeit im Vordergrund stehen, dass es um die Verbesserung der Maßnahme oder des Projektverlaufs geht und nicht um die Bewertung der Akteure.

5 Dokumentation und Bericht

Zum Abschluss der umgesetzten Maßnahmen und der dazugehörigen Evaluation mit Datenerhebung, -auswertung und -interpretation steht der Bericht. In diesem werden alle Schritte der Maßnahmenplanung und der Evaluation sowie die Ergebnisse und Effekte der Maßnahmen nachvollziehbar dokumentiert. In den folgenden Abschnitten sind Tipps und Hinweise für die Berichtlegung und Veröffentlichung der Evaluation zusammengestellt.

5.1 Ergebnisdarstellung

Es gibt verschiedene Ebenen von Evaluationsberichten, grob können zunächst interne und externe Berichte unterschieden werden. Für interne Berichte, zum Beispiel als Nachweis für die korrekte Verwendung von Fördergeldern, liegen oft klare Hinweise für die Gliederung und den Inhalt vor. In aller Regel werden diese internen Berichte nicht breit veröffentlicht.

Der Aufwand für einen externen Bericht lohnt sich vor allem für Evaluationen, die auch für andere interessant sein könnten. Nur durch diese Veröffentlichungen können weitere Interessengruppen, Entscheidungsträger, Kommunen, Unternehmen usw. von den Erfahrungen lernen und somit ein Wissens- und Technologietransfer ermöglicht werden.
Ein aussagekräftiger Bericht kann nach dem folgenden Schema angelegt werden:

- Problembeschreibung, Ist-Zustand
- Zielzustand auf allen Ebenen mit den dazugehörigen Indikatoren
- Zeit- und Maßnahmenplan
- Beschreibung aller Projektbeteiligten
- Beschreibung der Maßnahmenumsetzung und ggf. der Abweichungen
- Prozessevaluation
- Ergebnisse und Effekte einschließlich der Wirkungsevaluation
- Diskussion der Ergebnisse im Hinblick auf die gesetzten Ziele
- Zusammenfassung.

Grundsätzlich hilft bei der Erstellung, sich Gedanken über die sogenannten W-Fragen zu machen. Dazu gehören:

A. Was möchten Sie sagen?
Was genau ist der Inhalt des Berichts? Wird die Maßnahme mit den Effekten beschrieben? Geht es um den Prozess? Stehen die überwundenen Schwierigkeiten im Vordergrund oder der Nutzen für die Zielgruppen?

B. Weshalb möchten Sie es sagen?
Was ist das Ziel des Berichts? Dient er der internen Dokumentation oder dem Nachweis der Mittelverwendung? Soll er anderen als Anregung oder Entscheidungsgrundlage dienen? Soll er zu einem Wissens- oder Technologietransfer anregen? Soll er andere zum Handeln motivieren? Oder geht es um eher organisatorische und administrative Fragen?

C. Wer sind die Zielgruppen und welche Informationen benötigen sie?
An wen richtet sich der Bericht primär? Richtet er sich an die kommunale Verwaltungsspitze oder andere Entscheidungsträger? Oder adressiert er die Bevölkerung, die Öffentlichkeit, die Nutznießer des Vorhabens? Welche Informationen benötigen diese Zielgruppen und welches Vorwissen haben sie? Diese Entscheidungen haben maßgeblichen Einfluss darauf, ob der Bericht in der Fachsprache verfasst werden kann oder besser in allgemeinverständlicher Sprache sein sollte.

D. Wie möchten Sie den Bericht vermitteln?

Wie kann die Botschaft am besten kommuniziert werden? Ist eine schriftliche Information das Mittel der Wahl? Oder soll es einen interaktiven Dialog und Austausch geben?

PRINTMEDIEN	MULTIMEDIA	PERSÖNLICH
Newsletter	Website	Besprechungen
Faltblätter	CD/DVD	Seminare
Projektdatenblätter	Webbasierte Foren	Workshops
Pressemitteilungen	Telefonkonferenzen	Podiumsdiskussionen
...	Webinare	...

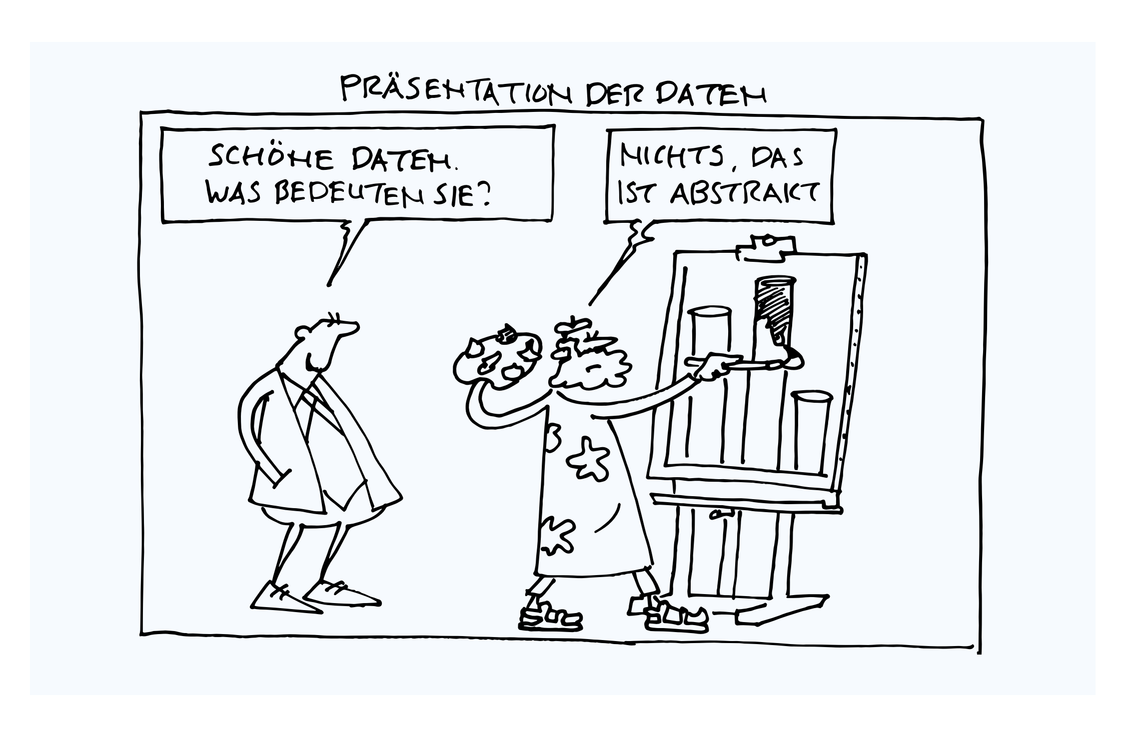
Tabelle 5-1: Vorschläge für Verbreitungskanäle

Die Verbreitungskanäle (siehe Tab. 5-1) sollten zu dem Projekt passen. Es können auch für unterschiedliche Zielgruppen mehrere Methoden kombiniert werden.



5.2 Datendarstellung

Der größte Teil eines Berichts wird aus Text bestehen. Tabellen, Diagramme und Grafiken lockern den Text nicht nur auf, sondern können einzelne Ergebnisse auch knapp und übersichtlich veranschaulichen. Dies gilt insbesondere für statistische Daten und Fakten oder für technische Informationen. Diese wären in Textform sehr unübersichtlich. Aber genau so, wie Texte zum Beispiel durch Verwendung von zu viel Fachsprache für Laien unverständlich oder schwer lesbar sein können, gibt es auch Grafiken, die mehr verwirren als aufklären.



Histogramme können helfen, Häufigkeiten darzustellen. Im Beispiel (Abb. 5-1) werden kategorisierte Häufigkeiten berichtet. In der Überschrift ist dokumentiert, wie viele Personen an der Befragung teilnahmen, in diesem Fall 813. Eine solche Angabe ist vor allem dann wichtig, wenn in der Grafik Prozentangaben berichtet werden, die Fallzahlen also dort nicht abgelesen werden können.

Die Häufigkeiten werden durch vertikale oder auch horizontale Balken oder Säulen dargestellt, daher auch die deutschen Begriffe Balken- oder Säulendiagramm. Für eine zusätzliche Veranschaulichung kann auch ein Polygon hinzugefügt werden. Dies wird durch Verbinden der oberen Mitte der Klassen (Säulen) durch gerade Linien erzeugt.

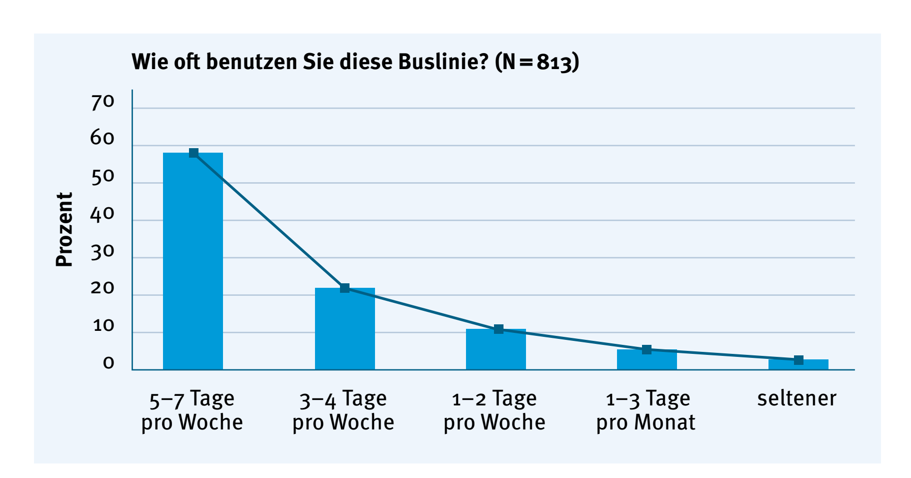


Abbildung 5-1: Beispiel für die Darstellung der Ergebnisse durch Histogramm (Säulendiagramm) und Polygon

Kreisdiagramme (auch Kuchen- oder Tortendiagramme genannt) werden für die bildliche Darstellung von Anteilen oder Größenverhältnissen verwendet. Sie sind dann gut lesbar, wenn die Stücke unterschiedlich farbig markiert werden. Es sollten nicht zu viele oder zu kleine Anteile sein, da sich das negativ auf die Lesbarkeit auswirkt. Im nachfolgenden Beispiel (Abb. 5-2) wird der jeweilige prozentuale Anteil von Frauen und Männern einer Stichprobe mit 813 Personen dargestellt.

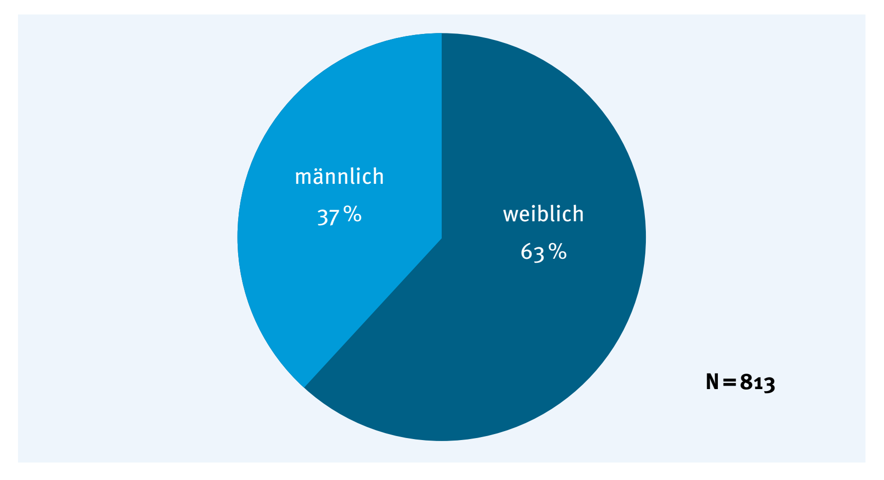


Abbildung 5-2: Beispiel für die Darstellung der Ergebnisse durch ein Kreisdiagramm

Bei der Verwendung eines *Streudiagramms* werden zwei Variablen gleichzeitig veranschaulicht (Abb. 5-3). Diese Art der Darstellung hilft dabei, eine statistische Korrelation zwischen zwei Variablen, also ein Zusammenhang sowie die Intensität darzustellen. Die Werte einer Variablen werden auf der horizontalen Achse (auch Abszisse oder x-Achse genannt), die andere Variable auf der vertikalen Achse (auch Ordinate oder y-Achse) eingetragen. Jede statistische Einheit (zum Beispiel jede befragte Person) erhält einen Punkt im Koordinatensystem. Üblicherweise werden auf der x-Achse die Werte eingetragen, welche die beeinflussenden oder erklärenden Faktoren sind, das sind die so genannten »unabhängigen Variablen«. Auf der y-Achse werden die »abhängigen« Variablen dargestellt, also diejenigen, die beeinflusst werden. Das folgende Beispiel stellt Radfahrten zur Schule pro Monat in Abhängigkeit vom Alter der Schülerinnen und Schüler dar (Abb. 5-3).

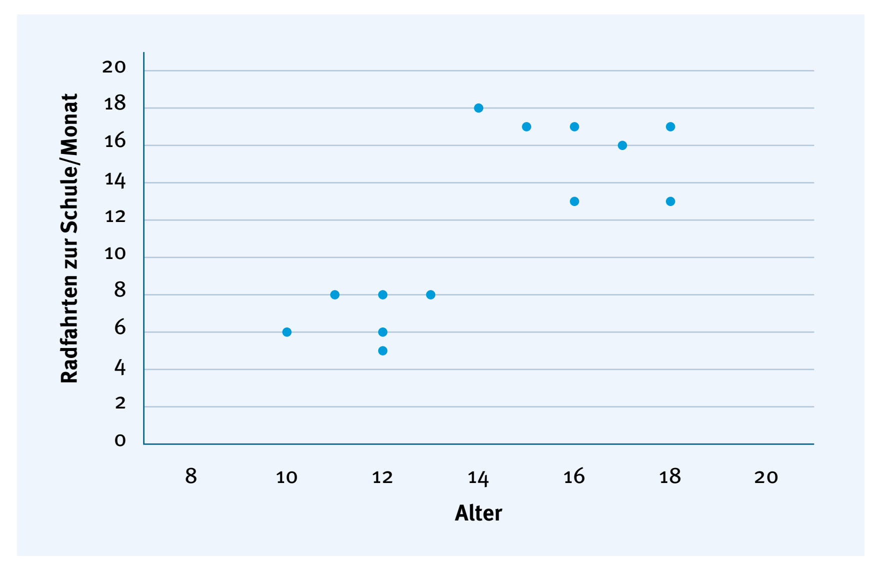


Abbildung 5-3: Beispiel für ein Streudiagramm

Entwicklungen über einen Zeitraum hinweg können anschaulich mit einem *Liniendiagramm* dargestellt werden. Auf der horizontalen Achse werden die Zeiteinheiten und auf der vertikalen Achse die Messergebnisse eingetragen. Im Zeitverlauf wird der prozentuale Anteil der Personenbeförderung im Öffentlichen Personennahverkehr in Evalstadt über mehrere Quartale hinweg angezeigt (Abb. 5-4).

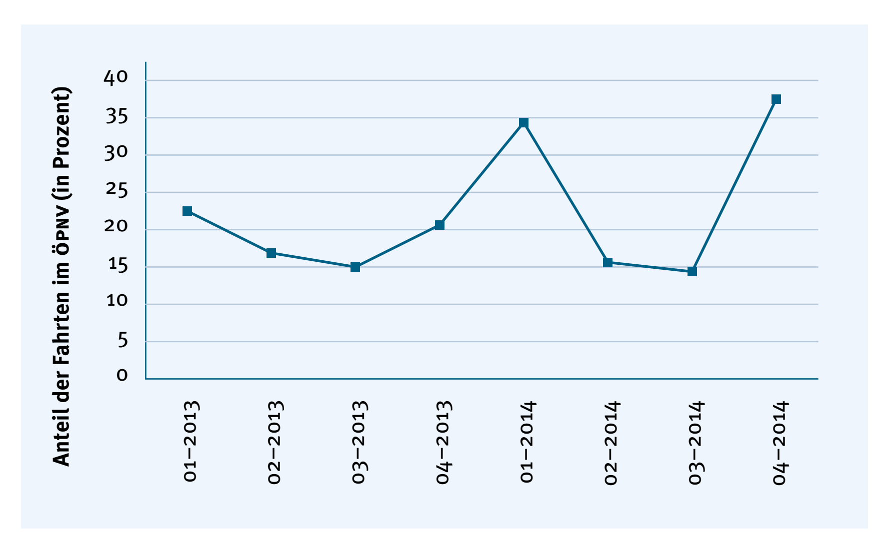


Abbildung 5-4: Beispiel für ein Liniendiagramm

Eine interessante grafische Darstellung einer Häufigkeitsverteilung ist das *Box-Plot-Diagramm* (deutsch: Kastengrafik). Es gibt einen schnellen und detaillierten Überblick darüber, wie sich die Werte verteilen. Der Kasten markiert alle Werte zwischen 25% und 75%, die Linien zeigen an, wie weit die beiden anderen Quartile streuen (Abb. 5-5). Die Linienendpunkte zeigen die Extremwerte auf. Kastengrafiken sind gut geeignet, um verschiedene Verteilungen zu vergleichen.

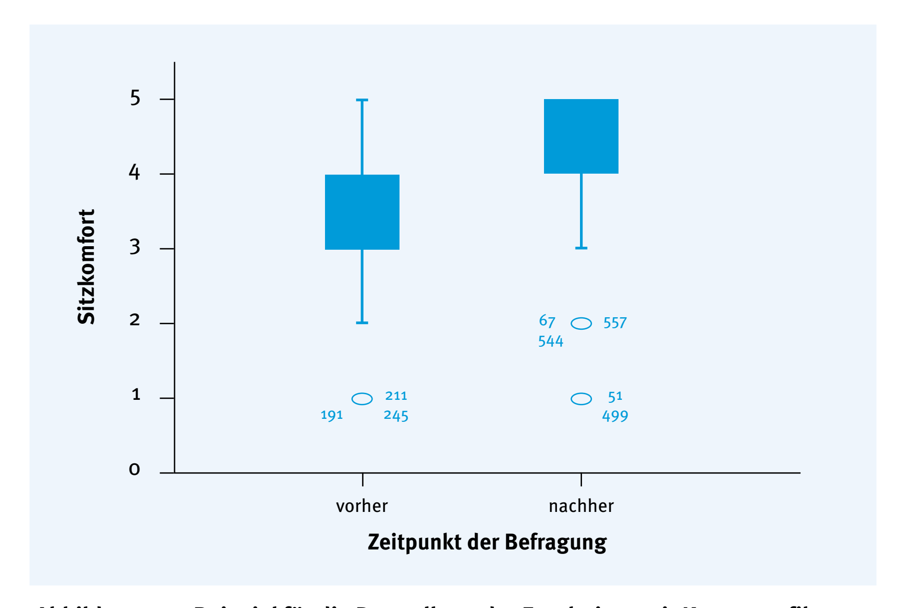


Abbildung 5-5: Beispiel für die Darstellung der Ergebnisse mit Kastengrafiken

Wie Daten nicht präsentiert werden sollten

»Vertraue keiner Statistik, die du nicht selbst gefälscht hast.« Dieses Zitat wird häufig in Statistikbüchern verwendet, um die Möglichkeit der Manipulation oder der bewussten oder unbewussten Fehlinterpretation zu unterstreichen.

Im folgenden Beispiel (Abb. 5-6) werden zwei Mal die gleichen Daten dargestellt. Es handelt sich um einen Vergleich des wahrgenommenen Sitzkomforts vor und nach einer durchgeführten Maßnahme. Der Unterschied in den Diagrammen geht nur auf die Wahl der Skala zurück. Im Beispiel links startet die y-Achse nicht bei Null, sondern erst bei 3,7 und die Abstände zwischen den Werten sind deutlich kleiner als auf der rechten Seite. Dadurch entsteht der Eindruck, dass im Projekt links ein größerer Zuwachs stattgefunden hätte.

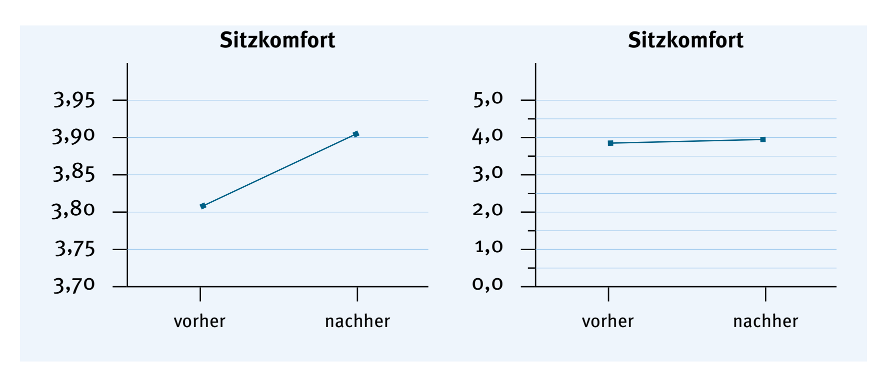


Abbildung 5-6: Darstellung einer y-Achsen-Verschiebung

Bei der Verwendung von Balken- oder Säulendiagrammen werden manchmal – auch durch Voreinstellungen in den Standardprogrammen – 3-D-Darstellungen verwendet. Sie sind fast immer schwerer zu lesen als zweidimensionale Grafiken. Jede neue Perspektive führt zu einem anderen Eindruck. Eine hohe Säule scheint von weit oben gesehen eher flach, während eine kurze Säule von unten gesehen schon recht groß aussehen kann (siehe Abbildung 5-7). Die Daten sind nur schlecht ablesbar und kaum vergleichbar.

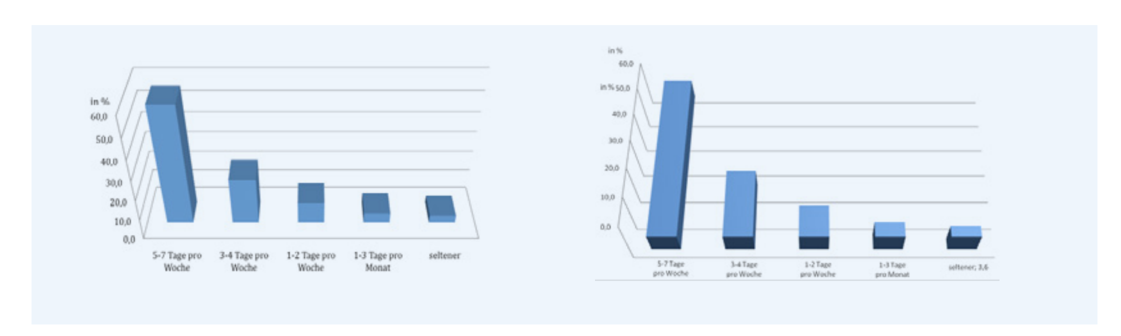


Abbildung 5-7: 3D-Grafik-Darstellung eines Histogramms

5.3 Übertragbarkeit

Erfolgreiche Projekte sollen Andere zum Nachahmen anregen. Aber wie schon mehrfach festgestellt, können auch die Erfahrungen aus dem Projektverlauf oder der Umsetzung, die weniger erfolgreich verliefen, sehr hilfreich und informativ sein. Aus Erfahrung wird man klug!

Die Übertragung einer Maßnahme oder eines Maßnahmenbündels von einer Stadt auf die andere ist weit verbreitet. Aber nur, weil sie in Evalstadt erfolgreich war, bedeutet dies nicht zwangsläufig, dass sie in jeder anderen Stadt ebenso erfolgreich umgesetzt werden kann.

Übertragbarkeit bezieht sich nicht einfach auf einzelne technische oder betriebliche Funktionen, sondern darauf, wie eine Maßnahme in den Stadtkontext passt und wie die Maßnahmen miteinander verknüpft sind, d. h. unter welchen Bedingungen und Kombinationen eine bestimmte Maßnahme oder ein Maßnahmenpaket andernorts mit einem vergleichbaren Erfolgsgrad angewendet werden kann.

Beurteilung der Übertragbarkeit

In der Praxis hängt die Übertragbarkeit zu einem großen Teil von den Eigenschaften der Maßnahmen selbst und den vorherrschenden Bedingungen für die Umsetzung in der Zielstadt ab.

Das bedeutet, dass in einem ersten Schritt alle zusammenhängenden Informationen, die zur Erklärung von Erfolg oder Misserfolg einer Maßnahme beitragen können, gesammelt und verbreitet werden müssen (z. B. als Bericht oder als Daten in einer Datenbank).

Kasten 14: Projektbeispiele aus europäischen Städten

Im europäischen Projekt CIVITAS wurden und werden Projektbeispiele und Maßnahmen, Berichte und Evaluationen zu thematischen Schwerpunkten zusammengefasst und veröffentlicht.
http://civitas.eu/mobility-solutions-page/
http://civitas.eu/civinet/civinet-german/

Die Schwerpunkte stellen erprobte Beispiele für den Öffentlichen Nahverkehr, die Sicherheit, die Bürgerbeteiligung, den Lieferverkehr und die Frachtlogistik, das Mobilitätsmanagement sowie für viele weitere Themen vor.

Der Erfolg einer bestimmten Maßnahme kann auch mit den Auswirkungen einer anderen Maßnahme in Zusammenhang stehen. Maßnahmenpakete führen oft zu Synergien mit sich gegenseitig verstärkenden Wirkungen.

Aus einer allgemeinen Sicht kann festgestellt werden, dass die wichtigsten Unterschiede bei der Übertragbarkeit von einer, der Ursprungsstadt, auf die andere Kommune, der Zielstadt, auf zwei wesentlichen Ebenen stattfinden. Das ist zum Einen die Ebene der lokalen Politik und der Verwaltung und die dort zu treffenden Entscheidungen und auf der anderen Ebene die öffentliche Wahrnehmung, Bewertung oder auch Akzeptanz.

An einer Maßnahmenübertragung sind immer zwei Akteure beteiligt – die Ursprungsstadt und die Zielstadt. Demzufolge gibt es auch zwei Blickwinkel, aus denen man das Thema Übertragbarkeit betrachten kann:

- Aus der Perspektive der Ursprungsstadt
- Aus der Perspektive der (potenziellen) Zielstadt.

Dementsprechend sind auch die Verantwortlichkeiten aufgeteilt. Die Ursprungsstadt liefert ihre Evaluierungsergebnisse und idealerweise noch ein paar Zusatzinformationen über die Rahmenbedingungen, unter denen die Maßnahme durchgeführt wurde. Diese dienen als Input für die ex-ante Evaluierung der Zielstadt, die die Sinnhaftigkeit einer Übertragung einschätzen muss.

Aufgaben der Ursprungsstadt

Damit sich andere Städte von erfolgreichen Maßnahmen inspirieren lassen können, geben wir für den Projekt-Bericht einige Empfehlungen.

1. Beschreiben Sie den Erfolg der Maßnahme in Ihrer Stadt am besten durch eine Wirkungsevaluation

Eine Voraussetzung für eine mögliche Übertragung ist, dass die Maßnahme bereits einmal mit Erfolg umgesetzt wurde. Was erfolgreich war, lässt sich am besten mit einer durchgeführten und gut dokumentierten Wirkungsevaluation darstellen.

2. Beschreiben Sie detailliert die Besonderheiten Ihrer Stadt

Wenn in Evalstadt der Anteil der Radfahrten auch deshalb so hoch ist, weil es innerorts praktisch keine Steigungen gibt und dort zudem sehr viele Studierende wohnen, so sind dies wichtige Information. Eine genaue Identifizierung der Besonderheiten der Stadt und ihrer Strukturen ist sehr hilfreich für die Beurteilung, ob eine Maßnahme auch in einer anderen Stadt funktionieren kann. Die Abbildung (5-8) zeigt einige relevante Aspekte, die bei der Beschreibung beachtet werden sollten.

Geographie	Akzeptanz	Institutionen
Topographie	Kultur	Rechtssystem
Bebaute Umgebung	Ökonomie	Politischer Rahmen
Natürliche Umgebung	Demographie	...
...	Soziale Muster	...
...	Durchsetzungsfragen	...

Abbildung 5-8: Kategorien zur Beschreibung der Besonderheiten in einer Stadt

3. Erstellen Sie Ursache-Wirkungs-Beziehungen zwischen den Auswirkungen der Maßnahmen und den Rahmenbedingungen der Stadt

Das Verständnis für die Ursache-Wirkungs-Beziehungen (siehe Kapitel 2.2 auf Seite 18) zwischen den Auswirkungen der Maßnahmen und den Besonderheiten der Stadt kann wesentlich für den Erfolg der Übertragung sein. Eine Gewichtung der beeinflussenden Faktoren hilft bei der Nachvollziehbarkeit.

4. Beschreiben Sie die gegenseitigen Abhängigkeiten zwischen den Maßnahmen

In vielen Städten ist es nicht die eine Maßnahme, die zum Erfolg einer nachhaltigen mobilen Stadt führt, sondern die gute Mischung sich gegenseitig ergänzender Projekte. Für eine Übertragbarkeit ist es wichtig zu wissen, wie die Maßnahmen zusammenhängen – egal ob räumlich, zeitlich, in ihrer Wirkung, in Bezug auf Zielgruppen und so weiter.

5. Beschreiben Sie die Umsetzung der Maßnahme so, dass andere diese nachahmen können

Eine gute Maßnahmenbeschreibung funktioniert ähnlich wie ein gutes Rezept: Es genügt nicht, nur ein schönes Foto vom fertigen Kuchen zu zeigen. Es braucht auch eine genaue Beschreibung der Zutaten einschließlich der Mengen, der Werkzeuge, einen zeitlichen Ablauf und die Beschreibung der einzelnen Handlungen. In Bezug auf Mobilitätsmaßnahmen wurden die meisten relevanten Aspekte im Kapitel Wirkungsevaluation beschrieben. Für einen Übertragbarkeitsbericht ergänzen Sie diese Punkte um die nötigen praktischen Hinweise und Tipps.

6. Veröffentlichen Sie Ihre Erkenntnisse

Neben klassischen gedruckten Ergebnisberichten oder Kurzbroschüren, die die wichtigsten Erkenntnisse zusammenfassen, bietet es sich auch an, Texte ins Internet zu stellen. Zum Beispiel auf der Homepage Ihrer Kommune oder Ihres Unternehmens. Zur weiteren Verbreitung ist es auch hilfreich, Ihre Ergebnisse in größeren Datenbanken eintragen zu lassen. So gibt es z. B. in Deutschland das Portal des Nationalen Radverkehrsplans, in dem sehr viele Ergebnisse zum Radverkehr gesammelt und aufbereitet werden (www.nationaler-radverkehrsplan.de). Im europäischen Kontext ist die Eltis Website eine gute Plattform für Evaluationsergebnisse kommunaler verkehrlicher Maßnahmen (www.eltis.org).

Die Aufgaben der Zielstadt

Die meisten Städte kennen ihre Fragestellungen, Rahmenbedingungen und lokalen Strukturen genau und suchen nach Anregungen für Maßnahmen. Es gibt viele gute Beispiele aus anderen Städten, einige Datenbanken für Beispiele stehen in Kasten 15. Wichtig ist, dass vor der Übertragung in die eigene Kommune anhand der Empfehlungen aus dem Abschnitt »Ursprungstadt« gründlich geprüft

wird, ob alle wichtigen Voraussetzungen und Erfolgsfaktoren für eine Übertragung gegeben sind oder geschaffen werden können.

Kasten 15: Wo finden Sie Fallstudien für urbane Mobilitätsmaßnahmen?

Ausgewählte Quellen für Fallstudien, Machbarkeitsstudien und Ideen für Maßnahmen zur Förderung einer nachhaltigen städtischen Mobilität finden Sie auf verschiedenen Websites, einschließlich:

- *www.civitas.eu*
- *www.eltis.org*
- *www.epomm.eu*
- *www.nationaler-radverkehrsplan.de*
- *www.mobilitymanagement.org*

Die dort vorgestellten Initiativen und Organisationen stellen sicher auch weitere Auskünfte und Informationen zur Verfügung.

6 Und jetzt: Zurück zur Realität

Für Evaluationen sprechen sehr viele und gute Argumente, dennoch werden Evaluationen in der Praxis selten durchgeführt. Dieses Buch soll anregen und das nötige Handwerkszeug liefern, mehr kommunale verkehrliche Maßnahmen zu evaluieren. In diesem abschließenden Kapitel werden gängige Vorurteile gegenüber Evaluationen auf ihren Wahrheitsgehalt hin untersucht und Empfehlungen zur Qualität und Gültigkeit einer Evaluation gegeben.

6.1 Einige Mythen über Evaluation

Es kursieren viele Vorurteile über Evaluationen. Die drei häufigsten sind: »viel zu kompliziert«, »zu teuer« und »Evaluationen dienen nur dazu, das Programm oder die Verantwortlichen zu kritisieren«.

»Evaluationen sind viel zu kompliziert«

Eine Evaluation kann tatsächlich eine sehr komplexe und langwierige Aufgabe sein und erfordert in der Planung und Durchführung Fachwissen und Erfahrung. Eine Evaluation besteht jedoch aus vielen kleineren und größeren Aufgaben, von denen manche in vielen Kommunen ohnehin standardmäßig durchgeführt werden. Wenn die einzelnen Teilschritte also zu Beginn der Planung in einem Gesamtkonzept strukturiert werden, wird es deutlich einfacher, die einzelnen Aufgaben anzugehen. Zudem können einzelne Teile auch an andere Abteilungen oder Projektpartner delegiert oder beauftragt werden.

»Evaluationen sind zu teuer«

Die Evaluationen, und hier besonders die Datenerhebung und Analyse, kosten Arbeitszeit und Geld. Wie schon bei der Argumentation zur Komplexität gilt aber auch bei diesem Argument, dass einige der Aufgaben ohnehin erfüllt werden müssen. Und eine Evaluation kann dazu beitragen, eine Kontrollfunktion auszuüben sowie eine möglichst hohe Kosteneffizienz und -effektivität der Maßnahmen zu sichern. Mit Hilfe von Evaluationen können frühzeitig Fehlentwicklungen oder unwirksame Maßnahmen korrigiert werden. Unnötige oder schwache Maßnahmen werden vermieden. Ein wichtiger Effekt der Evaluationen, das Lernen für zukünftige Projekte, kommt allerdings nicht der evaluierten Maßnahme direkt zugute. Der Effekt zeigt sich erst in späteren Projekten, er wird zu einem »öffentlichen Gut«. Im Rahmen von kommunalen Projekten bzw. mit öffentlichen Geldern finanzierten Projekten sollte dies jedoch selbstverständlich sein. Schließlich könnte noch die Option geprüft werden, ob eventuell die Evaluation teilweise oder ganz durch externe Fördermittel finanziert werden

kann, das gilt insbesondere für EU-Projekte. Gegebenenfalls können auch Partnerschaften mit Hochschulen oder Universitäten die nötige Unterstützung liefern.

»Evaluationen dienen nur dazu, das Projekt oder Programm zu bewerten und dann abzuschaffen oder die Verantwortlichen zu kritisieren.«

Eine Evaluation deckt durch den Einsatz verschiedener Methoden Schwachstellen auf. Die Umsetzung jeder Maßnahme birgt Risiken. Es muss also immer damit gerechnet werden, dass sich im Laufe eines Planverfahrens die Rahmenbedingungen oder Anforderungen so stark ändern, dass auch das Projekt angepasst werden muss. Das geschieht umso häufiger, je länger die Projekte laufen oder je aufwändiger sie sind. Je eher also durch eine Evaluation mögliche Schwachstellen erkannt werden, desto schneller und besser kann reagiert werden, um das Projekt oder die Maßnahme dennoch erfolgreich umzusetzen. Auf diese Weise wird das Risiko des Scheiterns – des Programms, des Projekts oder auch der Person – minimiert und zudem werden neue Handlungs- und Gestaltungsoptionen eröffnet. Die Evaluation fordert aber von allen Handelnden eine gewissen Offenheit und Kritikfähigkeit. Nur so kann die nötige Flexibilität aufgebracht werden, eine maßgeschneiderte und in der Wirkung optimale Maßnahme umzusetzen. Jedes starre Festhalten an einmal vor langer Zeit getroffenen Entscheidungen wäre nur schädlich und unprofessionell.



6.2 Störungen im Ablauf einer Evaluation

Im Laufe von Projekten ergeben sich immer wieder Situationen, die eine Evaluation im Ablauf stören oder sogar gefährden. Im Folgenden sollen die wichtigsten kurz skizziert werden.

Die Rahmenbedingungen für das Projekt verändern sich

Auf einmal ist die Finanzierung unsicher, die politischen Akteure setzen auf andere strategische Schwerpunkte oder Leitbilder, die gesellschaftliche Akzeptanz schwindet, die Notwendigkeit der Maßnahme nimmt ab, die rechtlichen Rahmenbedingungen ändern sich ...

Projekte sind praktisch täglich mit Veränderungen der Rahmenbedingungen konfrontiert. Manche kündigen sich länger an, manche kommen ganz unerwartet. Dies kann auch große Auswirkungen auf den Designplan der Evaluation haben. Muss die Maßnahme beispielsweise stark verändert werden, kann es passieren, dass die Daten der Vorher-Messung keinen Bezug mehr zu der »neuen« Maßnahme

haben. In einem solchen Fall kann versucht werden, andere Vergleichsdaten hinzuzuziehen. In jedem Fall sollten diese Veränderungen im Bericht dokumentiert werden.

Das Projekt verzögert sich

Ob durch die Veränderungen der Rahmenbedingungen oder aufgrund anderer Entwicklungen – Projektphasen und damit die Evaluationen sind immer in Gefahr deutlich länger zu laufen als ursprünglich geplant. Manchmal werden dann einzelne Phasen verkürzt, zum Beispiel die Testphase, um den Gesamtzeitplan einhalten zu können. Die Ergebnisse der Wirkungsevaluation des Pilotzeitraums können dann aufgrund der Verkürzung andere sein, als wenn die Phase wie ursprünglich geplant gelaufen wäre. Die Leistungsfähigkeit der Maßnahme kann in der Testphase dann nicht sicher bestimmt werden. Insbesondere bei Mobilitätsmaßnahmen spielen auch jahreszeitliche Effekte eine große Rolle. Projekte zur Förderung des Radverkehrs, die auf ihre Wirksamkeit hin überprüft werden sollen, werden im Winter andere Effekte zeigen als im Frühjahr. Hier wird deutlich, warum die parallel laufende Prozessevaluation so wichtig ist. Sie kann durch das Monitoring frühzeitig auf Verzögerungen hinweisen und gegebenenfalls die möglichen Folgen abschwächen.

Es werden zu viele Daten erhoben

Im Fachjargon wird der Effekt manchmal »Schrotschusstechnik« genannt. Es werden zunächst in alle Richtungen jede Menge Daten erhoben und gesammelt, um dann zu sehen »was dabei abfällt«. Das ist weder zeit- noch kosteneffizient und führt eher zu der Gefahr der Verzettelung. Dies kann die Evaluation aufwändiger machen, jedoch ohne einen Zugewinn an Erkenntnis.

Es werden zu wenige Daten erhoben

Nicht alle Datenerhebungen sind erfolgreich. Es kann z. B. sein, dass auf eine Umfrage deutlich weniger Menschen antworten, als es für die Stichprobe sinnvoll wäre, oder Daten sind nicht verfügbar, entweder weil sie doch nicht vorliegen oder nicht veröffentlicht und verbreitet werden dürfen. Gerade wenn die Datenerhebung an externe Dienstleistende vergeben wird, sollte mit Hilfe von Verträgen sichergestellt werden, wann, wer, wie und welche Daten liefert.

6.3 Die Qualität einer Evaluation erhöhen

Es gibt zwar viele Herausforderungen bei der Gestaltung und Durchführung einer hochwertigen Evaluation, aber es gibt auch bewährte Strategien, die die Qualität stärken und Wege aufzeigen, zukünftige Hindernisse zu überwinden.

Realistische Erwartungen formulieren

Nur durch die Evaluation alleine werden Arbeitsabläufe, Projekte und Maßnahmen nicht besser. Eine Evaluation macht also aus einem schwachen nicht automatisch ein starkes Projekt. Es bedarf auch der notwendigen Entscheidungen zur Anpassung.

Realistische Abläufe planen

Es ist nichts gewonnen, wenn für die Planung und Durchführung der Evaluation zu kurze Zeiträume zu Verfügung stehen. Veränderungsprozesse benötigen immer eine gewisse Zeit. Es muss z. B. auch möglich sein, Ergebnisse erst einmal »sacken« zu lassen, bevor sich die Beteiligten zu Veränderungen in einem Projekt entschließen können.

Rückmeldungen geben und einholen

Formelle und informelle Rückmeldungen helfen dabei, alle Projektbeteiligten »an Bord« zu halten. Wie geht es den Beteiligten mit der Evaluation? Sind sie mit den Abläufen zufrieden? Spüren sie den Nutzen der Evaluation oder empfinden sie diese als Last? Rückmeldungen sollten immer konstruktiv sein. Persönliche Bewertungen oder Kritik sollten sehr sparsam verteilt werden, und auch nur, wenn

die anderen dazu Bereitschaft signalisieren. Das Etablieren einer Art kollegialer Supervision »auf Augenhöhe« zum Beispiel im Rahmen regelmäßiger Treffen (jour fixe, Stammtisch) oder Video- und Telefonkonferenzen kann dabei helfen, die eigenen Lernerfahrungen während einer Evaluation zu verstehen und zu vertiefen.



Dauerhafte Speicherung der Daten sicherstellen

Viele Daten und Berichte sollten rund 10 Jahre lang aufgehoben werden. Bei besonders aufwändigen und langfristigen Projekten ergibt es Sinn, diese sogar noch deutlich länger aufzubewahren. Es muss also einerseits sichergestellt werden, dass genug (digitale) Speicherkapazität und Platz für die Sicherung zu Verfügung steht. Bei Datenerhebungen mit paper-pencil-Fragebögen können schnell viele Ordner zusammenkommen. Andererseits muss darauf geachtet werden, dass die Daten gemäß den Datenschutzrichtlinien sicher aufbewahrt werden und der Zugriff von Unbefugten vermieden wird (siehe Kapitel 9.4 auf Seite 79). Eine sorgfältige Dokumentation darüber, wo und wie welche Daten gelagert und gespeichert sind, erleichtern den späteren Zugriff enorm.

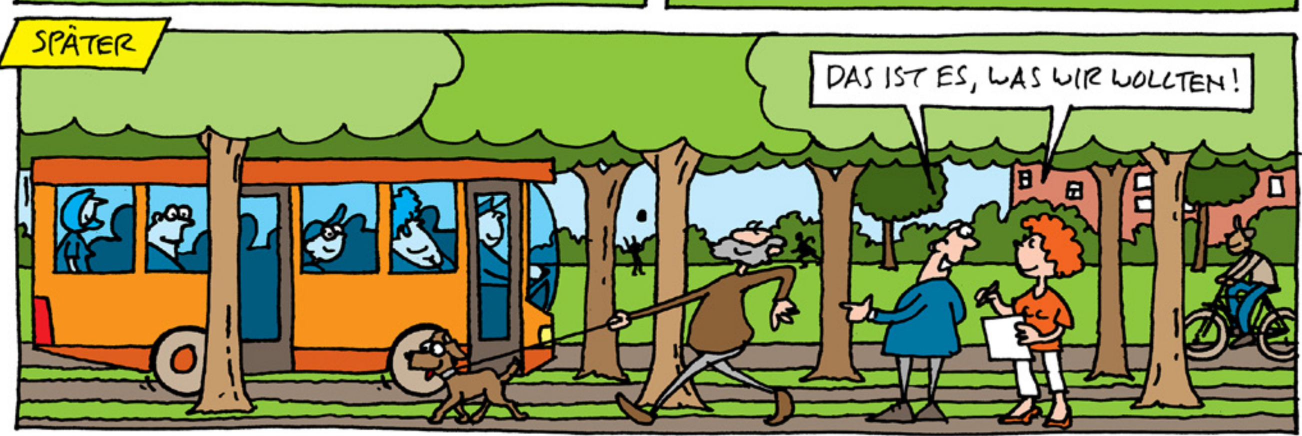

IST ES DAS, WAS WIR WOLLEN?
EVALUIEREN WIR DOCH MAL


ÜBERGEORDNETES ZIEL?
EINE SCHÖNE STADT FÜR DIE MENSCHEN


WAS HABEN WIR GEMACHT?
WIR HABEN AUTOS DURCH BUSSE ERSETZT


ERGEBNIS?
ZU VIELE BUSSE MIT ZU WENIGEN PASSAGIEREN


ALSO?
NICHT NUR BUSSE, SONDERN AUCH GEHEN UND RAD FAHREN


SPÄTER
DAS IST ES, WAS WIR WOLLTEN!

7 Das allerletzte Kapitel

Der Comic am Anfang des Buches und der auf der vorherigen Seite illustrieren, dass die scheinbar naheliegenden Lösungen für Mobilitäts- und Umweltprobleme nicht immer die richtigen sein müssen. Evaluation kann dabei helfen, dass Problem besser zu erkennen, geeignete Lösungen zu finden und sich damit für die passende, wirksame Maßnahme zu entscheiden. Durch Evaluation kann man viel Geld sparen, einerseits weil die Maßnahme nahezu optimal umgesetzt wird und andererseits, weil Fehlentwicklungen frühzeitig erkannt werden können und ein Gegensteuern möglich wird. Gute Datensammlungen sind eine verlässliche Entscheidungsgrundlage für kommunale Akteure und ermöglichen das Erkennen von Veränderungen und damit auch Wirkungen von zum Beispiel kommunalen verkehrlichen Maßnahmen. Fakten und harte Zahlen sprechen eine klare Sprache und jeder Politiker und jede Politikerin verwendet freudig positive Statistiken in der Kommunikation über den Erfolg einer Maßnahme. Aber es gilt auch aus Fehlern zu lernen. So sind ebenfalls scheinbar negative Ergebnisse der Evaluation sehr hilfreich, damit andere nicht die gleichen Fehler machen bzw. Erkenntnisse gewonnen werden können, wie man eine Maßnahme so verändert, dass sie positive Ergebnisse bringt. Transparenz ist dabei wichtig und Wählerinnen und Wähler wissen gern, wofür ihre Steuergelder eingesetzt werden.

Evaluation sollte ein integraler Bestandteil eines jeden Projektes sein. Sie sollte von Anfang an mitgedacht werden, z. B. welche Wirkungen werden von der Maßnahme erwartet, an Hand welcher Indikatoren kann man diese Wirkung messen, und wie sie mit einem ausreichenden Budget und Ressourcen ausstatten? Im Rahmen der CIVITAS Projekte wurden durchschnittlich etwa 10% der Projektmittel für Evaluation ausgegeben.

Dieses Buch soll das nötige Verständnis vermitteln, um kleinere Evaluierungsaufgaben selbst durchzuführen, aber auch um größere Projekte angemessen planen und begleiten zu können. Wenn finanziell möglich, ist eine externe Beauftragung der Evaluation an neutrale, fachkundige Dritte hilfreich. Eine enge Einbindung aller Projektbeteiligten ist jedoch mit zu planen und sowohl bei der Wirkungsevaluation als auch besonders bei der Prozessevaluation essentiell.

8 Glossar

Deutsch	Englisch	Definition
Abbruchrate	dropout rate	Beschreibt die Quote von Untersuchungsteilnehmenden, die nicht bis zum Ende der Maßnahme teilgenommen haben
Bedarfsanalyse	demand analysis	Eine Analyse des Bedarf an bestimmten Dienstleistungen oder Verkehrsmitteln und Verbindungen in einem bestimmten Gebiet
Datenanalyse	data analysis	Statistische Auswertung der erhobenen Daten
Dateninterpretation	data interpretation	In diesem Schritt wird entschieden, ob die angenommene Hypothese durch die erhobenen Daten bestätigt oder abgelehnt werden können.
Evaluationsdesign	evaluation design	Beschreibt detailliert, wie die Evaluation aufgebaut ist und sollte Fragen zu Zielen, Methoden und möglichen Ergebnissen beantworten

Experimentalgruppe	experimental group	Versuchsgruppe, bei der die Versuchsvariable manipuliert wird. Die erhobenen Daten werden mit der einer zweiten Gruppe (Vergleichs- oder Kontrollgruppe) verglichen
Goldstandard	gold standard	Den Goldstandard erfüllen Evaluationen, die mit einem randomisierten Kontrollgruppen-Design mit Vorher-Nachher-Messung arbeiten
Indikator	indicator	Eine Messgröße die zeigt, ob eine bestimmte Bedingung erfüllt wird oder ob bestimmte Ziele erreicht bzw. nicht erreicht wurden
Kategorien	categories	Werden gebildet, um qualitative Daten möglichst eindeutig zu gruppieren, das hilft bei der Interpretation
Kontrollgruppe	control group	Gruppe von Versuchsteilnehmenden, bei der die Versuchsvariable nicht manipuliert wird und die zufällig dieser Gruppe zugeordnet wurden. Die erhobenen Daten werden mit denen der Experimentalgruppe verglichen
Management-Zusammenfassung	management summary	Eine an Führungskräfte gerichtete Zusammenfassung der Ziele, Methoden und Ergebnisse der Evaluation
Maßnahme	measure	Ein Eingriff, der eine gewünschte Veränderung bewirken soll
Mittelwert	mean	Kennzeichnet die zentrale Lage einer Verteilung. Je nach Skalenniveau gibt es verschiedene Möglichkeiten, Mittelwerte zu berechnen
Modal Split	modal split	Beschreibt die Anteile verschiedener Verkehrsmittel am gesamten Verkehrsaufkommen in einem bestimmten Gebiet oder für einen bestimmten Weg
Monitoring	monitoring	Überbegriff für alle Arten der Erfassung oder Überwachung eines Vorgangs oder Prozesses mit Hilfe geeigneter Beobachtungssysteme
Panel	panel	Eine feststehende Gruppe von Personen, die über mehrere Befragungen hinweg an einer Evaluation oder Studie beteiligt werden
Panel-Design	panel design	Design, bei dem dieselbe Personengruppe (Stichprobe) zu unterschiedlichen Zeitpunkten befragt oder untersucht wird. So lassen sich sowohl Veränderungen zwischen den verschiedenen Versuchspersonen als auch innerhalb der gleichen Versuchspersonen feststellen

Population (Grundgesamtheit)	target population	Definition einer Menge von Personen, für die die Aussagen gelten sollen
Primärdaten	primary data	Werden unmittelbar erhoben, z. B. durch einen eigens entwickelten Fragebogen
Prozessevaluation	process evaluation	Beurteilt kontinuierlich und begleitend den Ablauf der Maßnahmenplanung und -umsetzung, untersucht Treiber und Widerstände
Qualitative Daten	qualitative data	Daten, die ein empirisches Phänomen beschreiben oder charakterisieren, oft in schriftlicher oder audio-visueller Form, sie ermöglichen das Verstehen von Verhalten und Zusammenhängen, basierend auf subjektiver Interpretation
Quantitative Daten	quantitative data	Daten, die ein empirisches Phänomen numerisch definieren, sie ermöglichen das Verstehen von Verhalten und Zusammenhängen, basierend auf objektiver Interpretation
Quasi-experimentell	quasi-experimental	Design, bei dem die Versuchspersonen nicht per Zufall, sondern nach bestehenden Merkmalen (z. B. Wohnort) der Experimental- oder der Vergleichsgruppe zugeordnet werden
Randomisieren	randomisation	Die zufällige Zuordnung der Versuchspersonen zur Experimental- bzw. Kontrollgruppe
Referenzsituation	reference scenario	Hypothetischer Vergleichsfall, in dem so weiter gemacht wird wie bisher.
Sekundärdaten	secondary data	Daten, die während der Evaluation nicht selbst erhoben werden, sondern aus anderen Erhebungen oder Quellen vorliegen
Skala	scale	Zuordnungsvorschrift bei der Objektivierung einer abgestuften Bewertung, die Zuordnung von Zahlen zu Objekten oder Ereignissen müssen strukturtreue Abbildungen sein
Standardabweichung	standard deviation	Maß für die Streuung von Daten in der (deskriptiven) Statistik, eine Methode, um Abweichungen vom Durchschnitt zu messen und zu vergleichen
Tiefeninterview	in-depth interview	Ein offenes oder teil-strukturiertes Interview, welches durch die Art der Fragen auch bislang nicht bewusstes Wissen, Erfahrungen oder Erleben zu erfassen versucht
Triangulation	triangulation	Die Kombination verschiedener Methoden, um eine Maßnahme umfassend bewerten zu können

Ursache-Wirkungskette	cause-effect chain	Methode für die Abschätzung der Wirkung einer Maßnahme, bei der Kausalzusammenhänge verdeutlicht werden sollen
Vorher-Nachher-Vergleich	before and after comparison	Methode zur Bestimmung von Kausalzusammenhängen durch Datenerhebung vor und nach der Maßnahme
Wirkungsevaluation	impact evaluation	Wirksamkeitsprüfung einer spezifischen Maßnahme; soll Aufschluss geben über alle positiven wie negativen Wirkungen einer Maßnahme und möglichst Kausalzusammenhänge verdeutlichen
Zeitreihenanalyse	time series analysis	Analyse, bei der die Variablen zu mehreren Zeitpunkten gemessen werden; so lassen sich sowohl Veränderungen zwischen den Versuchspersonen als auch innerhalb der Versuchspersonen feststellen
Zielgruppenanalyse	target group analysis	Methode für die Abschätzung der Wirkung einer Maßnahme, bei der alle relevanten Personen oder Gruppen separat betrachtet werden

9 Literatur und Internetseiten

9.1 Kommentierte Literaturliste zu Evaluationen

BORTZ, JÜRGEN; DÖRING, NICOLA (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler.* 4. überarbeitete Auflage, Heidelberg: Springer.
Das Buch ist *das* Standardwerk für die Sozialwissenschaften und daher unbedingt zu empfehlen. Es eignet sich allerdings nicht für Einsteiger.

FLICK, UWE (Hg.) (2006). *Qualitative Evaluationsforschung: Konzepte – Methoden – Umsetzungen.* Reinbek: Rowohlt.
Das Buch verfolgt die Zielstellung einer »mehr oder minder systematischen Einführung«, die Bestandsaufnahme mit einem Beitrag zur konzeptionellen Weiterentwicklung qualitativer Evaluation. Dabei werden konkrete (offene) Erhebungsverfahren vorgestellt (Gruppendiskussionen, Interviews, Vignetten) und beispielhafte Evaluations- bzw. Forschungsprojekte vorgestellt. Insgesamt wird eine breite Palette von Anwendungsfeldern von der Kinder- und Jugendhilfe über Organisationsentwicklungsprozesse, Umweltpolitik, Hochschule, Wissenschaft und Entwicklungszusammenarbeit dargestellt.

Forschungsgesellschaft für Straßen und Verkehrswesen. Arbeitsgruppe Verkehrsplanung. (Hg.) (2012). *Hinweise zur Evaluation von verkehrsbezogenen Maßnahmen.* Köln: FGSV Verlag GmbH.
Eine kurz gehaltene Übersicht über die Kernpunkte der Evaluation von verkehrlichen Maßnahmen für Personen mit Vorwissen in der quantitativen und qualitativen Datenerhebung und -auswertung.

FÜRST, DIETRICH; SCHOLLES, FRANK (Hg.) (2007). *Handbuch Theorien und Methoden der Raum- und Umweltplanung.* Detmold: Rohn Verlag.

Das Lehrbuch wendet sich an Studierende der Planungswissenschaften. Insbesondere der zweite Teil des Bandes ist für eine Evaluation interessant, hier werden Methoden der Zielformulierung und Bewertung (z. B. Präferenzmatrix, Relevanzbaum, Nutzwertanalyse, SWOT-Analyse etc.) für die Anwendungsgebiete der Raum- und Umweltplanung knapp und verständlich dargestellt.

GOLLWITZER, MARIO; JÄGER, REINHOLD S. (2009). *Evaluation Kompakt*. Weinheim [u. a.]: Beltz PVU.
Das Buch stellt anwendungsnah Einsatzbereiche und Methodik von Evaluation vor, sowohl für psychologische als auch für nicht-psychologische Anwendungsbereiche.

GÖTZE, WOLFGANG; DEUTSCHMANN, CHRISTEL; LINK, HEIKE (2002). *Statistik: Lehr- und Übungsbuch mit Beispielen aus der Tourismus- und Verkehrswirtschaft*. München: Oldenbourg Verlag.
Das Lehrbuch wendet sich an Studierende interdisziplinärer Studiengänge wie der Tourismus- und Verkehrswirtschaft, der Wirtschaftsinformatik und des Wirtschaftsingenieurwesens. Mathematische Vorkenntnisse sind jedoch nicht zwingend erforderlich. Neben Statistik wird im Buch auch die Datenerhebung besprochen.

KUCKARTZ, UDO; DRESING, THORSTEN; RÄDIKER, STEFAN; STEFER, CLAUS (2008). *Qualitative Evaluation: Der Einstieg in die Praxis*. 2. Auflage, Wiesbaden: VS Verlag.
Am Beispiel einer von den Autoren realisierten Evaluation führt das Buch in die Planung und praktische Durchführung von qualitativer Evaluation ein. Anhand einer Schritt-für-Schritt-Anleitung werden die Bestandteile eines Evaluationsprozesses von der Gegenstandsbestimmung, Erhebung und Codierung der Daten bis hin zur computergestützten, kategorienbasierten Auswertung mit Hilfe der Software MAXQDA sowie der Berichtlegung dargestellt.

KUCKARTZ, UDO; EBERT, THOMAS; RÄDIKER, STEFAN; STEFER, CLAUS (2009). *Evaluation online: Internetgestützte Befragung in der Praxis*. Wiesbaden: VS Verlag.
Dieses Buch gibt eine detaillierte und praxisnahe Einführung in die Konzeption, Durchführung und Auswertung von Online-Befragungen mit verschiedenen Methoden.

WIDMER, THOMAS; BEYWL, WOLFGANG; FABIAN, CARLO (HG.) (2009). *Evaluation. Ein systematisches Handbuch*. Wiesbaden: VS Verlag.
Ein systematischer Vergleich der Evaluationspraxis in Deutschland, Österreich und der Schweiz in verschiedenen Themenbereichen wie Bildung, Energie- und Umweltpolitik, Gesundheit und Raumentwicklung.

9.2 Webseiten

Auf der Webseite der Länderinitiative Kernindikatoren (LIKI) *http://www.lanuv.nrw.de/liki-newsletter/index.php?mode=home* werden verschiedene umweltspezifische Nachhaltigkeitsindikatoren aufgelistet, inklusive deren Bedeutung und etwaigen Zielstellungen. Diese werden regelmäßig aktualisiert. Man kann dort z. B. vergleichen, wie der Jahresmittelwert bei der Luftqualität des jeweiligen Bundeslandes liegt und welche Ziele angestrebt werden.

Centrum für Evaluation (CEval) der Universität des Saarlandes: *http://www.ceval.de*
Das Centrum für Evaluation (CEval) ist ein wissenschaftliches Institut der Universität des Saarlandes, das grundlagen- und anwendungsorientierte Forschung im Bereich der Evaluation betreibt, Evaluationen aller Art durchführt und Beratungsleistungen sowie Aus- und Weiterbildungsmaßnahmen in den Themenfeldern Monitoring und Evaluation anbietet. Über die Internetseite werden sehr viele Publikationen zur Verfügung gestellt.

CIVITAS: *http://civitas.eu/mobility-solutions-page*
In europäischen CIVITAS-Projekten wurden Berichte und Evaluationen von Beispielen kommunaler verkehrlicher Maßnahmen zu thematischen Schwerpunkten zusammengefasst und veröffentlicht. Die Schwerpunkte stellen erprobte Beispiele für den Öffentlichen Nahverkehr, die Sicherheit, die Bürgerbeteiligung, den Lieferverkehr und die Frachtlogistik, das Mobilitätsmanagement sowie für viele weitere Themen vor.

European Commission (2008). EVALSED – The resource for the evaluation of socioeconomic development: *http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf*

Evalsed besteht aus zwei Teilbereichen: THE GUIDE will besonders politische Entscheidungsträger bei der Beauftragung und Interpretation von Evaluationsstudien unterstützen, The SOURCEBOOK gibt praktische Tipps für die Durchführung von Evaluationen.

Forum Qualitative Sozialforschung: *http://www.qualitative-research.net/index.php/fqs/index*
Fachartikel aus dem Methodenspektrum der qualitativen und quantitativen Forschung in Deutsch, Englisch und Spanisch

IFRC Monitoring and Evaluation web page (2011): *http://www.ifrc.org/MandE*
Auf den Seiten der International Federation of Red Cross and Red Crescent Moon (IFRC) finden sich ein guter Einsteigerleitfaden zu den Themen Monitoring und Evaluation sowie weitere hilfreiche Ressourcen. Die Webseite und Materialien sind allesamt in englischer Sprache.

MaxSumo (2009) Guidance on how to plan, monitor and evaluate mobility projects; *http://www.mobilitymanagement.org/index.phtml?Main_ID=2174&ID1=2359&id=2359*
Diese Seite gibt praktische Tipps für die Planung, Umsetzung und Evaluation von Maßnahmen im Bereich des Mobilitätsmanagements.

Gesundheitsökonomische Bewertungsinstrumente (HEAT) für Zufußgehen und Radfahren: Methodik und Benutzerhandbuch. Ökonomische Bewertung von Verkehrsinfrastruktur und Verkehrspolitik *http://www.euro.who.int/de/health-topics/environment-and-health/Transport-and-health/publications/2011/health-economic-assessment-tools-heat-for-walking-and-for-cycling.-methodology-and-user-guide.-economic-assessment-of-transport-infrastructure-and-policies.-2014-update*
Diese Seite und der herunterladbare Bericht beschreibt die Methodik von HEAT (Health Economic Assessment Tool, entwickelt von der WHO und weiteren Partnern), mit dessen Hilfe gesundheitsökonomische und ökologische Effekte von Maßnahmen, die das Zufußgehen und Radfahren fördern, abgeschätzt werden können.

9.3 Kommentierte Literaturliste zu Datenerhebung und Analyse

BORTZ, JÜRGEN; SCHUSTER, CHRISTOF (2010). *Statistik für Human- und Sozialwissenschaftler.* Lehrbuch mit Online-Materialien, 7. Aufl. Berlin: Springer.
Ein Klassiker für den sozialwissenschaftlichen Bereich. Sehr ausführlich, allerdings setzt es mathematische Vorkenntnisse voraus und für die Evaluation ist nur ein kleiner Teilbereich interessant.

DULLER, CHRISTINE (2007). *Einführung in die Statistik mit EXCEL und SPSS. Ein anwendungsorientiertes Lehr- und Arbeitsbuch.* 2. überarbeitete Auflage, Heidelberg: Physica.
Das Buch bietet eine leicht verständliche Einführung in die beschreibende und schließende Statistik. Die Methoden der Statistik werden in EXCEL und SPSS umgesetzt. Zahlreiche Beispiele sind mit vollständigem Lösungsweg angeführt. Weitere Übungsbeispiele mit Lösungen ergänzen die Darstellung, daher ist dieses Buch auch gut zum Selbststudium geeignet. Lediglich die gezeigte SPSS-Version ist etwas veraltet.

HELLFERICH, CORNELIA (2011). *Die Qualität qualitativer Daten: Manual für die Durchführung qualitativer Interviews.* Wiesbaden: Verlag für Sozialwissenschaften.
Der Band ist ein praktisches Lehrbuch für die Konstruktion von qualitativen Leitfadeninterviews. Es gibt konkrete Hilfestellung zu den einzelnen Schritten der Leitfadenkonstruktion und schlägt pragmatische Wege zur Durchführung vor. Musterformulare und -briefe – auch zum Datenschutz – runden das Werk ab.

JANSSEN, JÜRGEN; LAATZ, WILFRIED (2013). *Statistische Datenanalyse mit SPSS. Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests.* 8. Auflage. Berlin: Springer.
Der Titel ist Programm, das Buch bietet eine gute, anwendungsbezogene Einführung in die Software.

MANDERSCHEID, KATHARINA (2012). *Sozialwissenschaftliche Datenanalyse mit R.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Wer mit der freien Software R arbeitet, hat mit diesem sehr kostengünstigen Buch ein nützliches Nachschlagewerk, aber auch eine Art Lehrbuch, das Schritt für Schritt durch die Datenauswertung führt. Das Buch richtet sich explizit an SozialwissenschaftlerInnen und deren spezifische Daten und Analysetechniken.

RUMSEY, DEBORAH J. (2010). *Statistik für Dummies*. 2. überarbeitete Auflage. Indianapolis: Wiley Publishing.
Dieses Buch zeigt, dass die Auswertung und Interpretation von quantitativen Daten Spaß machen kann.

SCHNELL, RAINER; HILL, PAUL B.; ESSER, ELKE (2013). *Methoden der empirischen Sozialforschung*. München: Oldenbourg Verlag.
Der Band beschreibt die Erhebungsmethoden des quantitativen Forschungsstils und geht auf die wissenschafts- und messtheoretischen Hintergründe ein. Das Lehrbuch gibt einen prägnanten und fundierten Überblick über Hypothesengenerierung, Stichprobenplan etc.

Software für Online-Erhebungen

Inzwischen gibt es ein vielfältiges Angebot von Softwaretools, um Onlineerhebungen zu entwerfen und diese für die Umfrage zu veröffentlichen. Einen generellen Überblick bietet die folgende Seite: *http://www.gesis.org/unser-angebot/studien-planen/online-umfragen/software-fuer-online-befragungen/freie-software-open-source/*

Weitere Anbieter:
esurveys.pro: *http://www.esurveyspro.com/*
LimeSurvey: *http://www.limesurvey.org/*
NetQ: *http://www.netq-umfrage.de/*
QuestionPro: *http://www.questionpro.com/de/*
SoCI Survey: *https://www.soscisurvey.de/*
survey monkey: *https://de.surveymonkey.com/*
Surveysystem: *http://www.surveysystem.com/sscalc.htm*

Software für die Datenauswertung

Statistische Auswertung und Analyse mit Microsoft Excel
Nicht zu unterschätzen ist das gängige Tabellenkalkulationsprogramm Excel von Microsoft. Hier sind viele statistische Auswertungen möglich, die auch über die deskriptive Statistik hinausgehen wie statistische Test, Korrelations- und Regressionsberechnungen. Im Unterscheid zu den gängigen Statistikpakten müssen die Berechnungsformeln in der entsprechenden Kommandozeile aufgebaut werden. *http://www.microsoft.com*

Auch das Datenbankprogramm ACCESS von Microsoft ist für die Durchführung einer quantitativen Analyse interessant. Mit der Formularansicht kann z. B. ein Fragebogen für die Dateneingabe konstruiert werden, der auch mögliche Falscheingaben kontrolliert. *http://www.microsoft.com*

Statistische Auswertung und Analyse mit IBM® SPSS® Statistics
http://www-01.ibm.com/software/de/analytics/spss/
PASW Statistics (früherer Name SPSS) ist ein professionelles Datenanalyseprogramm, das u. a. im wissenschaftlichen Kontext eine große Verbreitung hat. Es richtet sich vor allem an geschulte Anwender mit Vorkenntnissen in der Datenanalyse. Es gibt eine kostenlose Testversion für 21 Tage.

Statistische Auswertung und Analyse mit SAS/STAT® Software
http://www.sas.com/offices/europe/germany/software/analytics/index.html
Die Software richtet sich in erster Linie an Fachanwender. Laut Firmenangaben ist SAS Analytics die weltweit mit Abstand leistungsfähigste Analyseumgebung.

Statistische Auswertung und Analyse mit R: *http://www.r-project.org/*

»R« ist eine freie Programmiersprache für statistisches Rechnen und statistische Grafiken. »R« läuft in einer Kommandozeilenumgebung und ist daher nur für Geübte einsetzbar. Die Software kann kostenlos von der Internetseite herunter geladen werden. Hinweise für die Benutzung bietet u. a.: MANDERSCHEID, KATHARINA (2012). *Sozialwissenschaftliche Datenanalyse mit R*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Khan Academy. *http://www.khanacademy.org/*
Kostenlose Sammlung von kurzen Tutorials und Videos (in englischer Sprache)

Stat Trek. *http://stattrek.com/*
Kostenlose Sammlung von kurzen Tutorials und Videos zu Statistik und Algebra (in englischer Sprache)

Statistika ist ein kommerzielles Datenanalyseprogramm der Firma StatSoft. Auch hier können alle wesentlichen statistischen Analysemöglichkeiten für eine quantitative Auswertung genutzt werden. *http://www.statsoft.de/*

Die Softwarefirma Stata Corp. bietet ein kommerzielles Statistikprogramm an, das komplexe Datensätze verarbeiten kann. Auch dieses Programm ist über Menüführung oder eine Syntaxprogrammierung bedienbar. *http://www.stata.com/*

Software für qualitative Datenauswertung

Atlas Ti ist eine der gängigsten Softwarelösungen für die qualitative Datenauswertung. Die Software wurde von Thomas Muhr an der TU Berlin entwickelt und ist aus dem Ansatz der Grounded Theory entstanden. Auswertung von Bild- und Filmmaterial ist mit dieser Software ebenfalls möglich. *http://www.atlasti.com/de/index.html*

MAXQDA wurde von Udo Kuckartz an der Philipps Universität Marburg entwickelt und ist aus dem Ansatz der qualitativen Inhaltsanalyse nach Mayring entstanden. Auswertung von Bild- und Filmmaterial ist mit dieser Software ebenfalls möglich. *http://www.maxqda.de/*

9.4 Datenschutz

Der Kontakt zu Datenschutzbeauftragten der Kommunen oder auch Betrieben sollte bei einer durchzuführenden Evaluation immer gesucht werden. Die Datenschutzbeauftragten können die geplanten qualitativen oder quantitativen Erhebungen hinsichtlich datenschutzrechtlicher Kriterien überprüfen. Evaluatoren sind immer auf der sicheren Seite wenn von dieser Stelle ein datenschutzrechtliches »Unbedenklich« attestiert wird.

Ein Orientierungspunkt für die Verwahrung und Sicherung von Primärdaten sollte die Broschüre der Deutschen Forschungsgemeinschaft sein, die in ihrer Empfehlung nützliche Anweisungen beschreibt: *http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf*

Für den Bereich des Datenschutzes von qualitativen Evaluationen gibt der Band »Die Qualität qualitativer Daten: Manual für die Durchführung qualitativer Interviews« von Cornelia Hellferich sehr gute Hinweise. Der Band enthält Mustererklärungen zum Einverständnis von Probanden und entsprechende Datenschutzerklärungen (siehe Kapitel 9.3 auf Seite 77).

Für den quantitativen Bereich gibt das Lehrbuch »Methoden der empirischen Sozialforschung« der Autoren Rainer Schnell, Paul B. Hill und Elke Esser konkrete Formulierungshilfen für die Gewährleistung des Datenschutzes insbesondere bei standardisierten Fragebögen (siehe Kapitel 9.3 auf Seite 77).

9.5 Allgemeine Literatur

OECD (1997). *Proceedings Towards sustainable transportation*. The Vancouver conference, Vancouver 24.–27. March 1996, Vancouver

SCHOLL, A. (2011). Motiviert durch Erfolg und Misserfolg. wissens.blitz (17). *http://www.wissensdialoge.de/erfolg-und-misserfolg*

Zeitschrift für Evaluation (ZfEv), *http://www.zfev.de*. Die halbjährlich erscheinende Zeitschrift behandelt theoretische, methodische und konzeptionelle Fragen, stellt Anwendungsbeispiele vor und diskutiert Fragen.

10 Vorhandenes Datenmaterial und Statistiken

Es gibt eine Fülle von verfügbaren Daten, die sich aber lokal in ihrer Aufbereitung unterscheiden. Zu den relevanten Sekundärdaten für mobilitätsbezogene Evaluationen gehören Einwohner- und Beschäftigungsdaten, Umweltdaten, Kfz-Bestandsdaten, Verkehrsdaten und Verhaltensdaten.

10.1 Einwohner- und Beschäftigungsdaten

Demographische Daten auf Bundes- oder Länderebene bieten das Statistische Bundesamt *(http://www.destatis.de/)* und die Statistischen Landesämter. Auch die Einwohnermeldeämter erheben entsprechende Informationen.

Der Deutsche Städtetag veröffentlicht seit 1890 jährlich das Statistische Jahrbuch Deutscher Gemeinden. *(http://www.staedtetag.de/dst/inter/publikationen/statistikjahrbuch/index.html)*

Informationen zu Beschäftigtenstatistiken, den Berufen und den Arbeitsmarkt liefern die Statistiken der Bundesagentur für Arbeit *(http://www.statistik.arbeitsagentur.de/)* sowie das Institut für Arbeitsmarkt- und Berufsforschung *(http://www.iab.de)*. Viele Städte und Gemeinden führen darüber hinaus eigene Statistiken in dem Themenfeld.

10.2 Umweltdaten

Luftschadstoffe

Allgemeine Informationen zu Luftschadstoffen aus dem Luftmessnetz gibt es auf der Seite des Umweltbundesamtes. Dort finden sich auch Links zu den Informationsseiten der einzelnen Bundesländer: *http://www.umweltbundesamt.de/themen/luft/messenbeobachtenueberwachen/luftmessnetze-der-bundeslaender*

Das Umweltbundesamt stellt auch eine Übersichtskarte zu den Luftqualitätsdaten bereit. Es ist möglich, nach Tag und Schadstoff auszuwählen und für jede einzelne Station Dateien mit den Messwerten runterzuladen: *http://www.umweltbundesamt.de/daten/luftbelastung/aktuelle-luftdaten*

Lärm

Das Umweltbundesamt stellt umfangreiche Informationen zur Umgebungslärmrichtlinie und zur Lärmkartierung allgemein zur Verfügung: *http://www.umweltbundesamt.de/themen/verkehr-laerm/umgebungslaermrichtlinie*

Unter den Ergebnissen der Lärmkartierung gibt es eine Tabelle mit einer Übersicht zur Belastung in einigen Städten: *http://www.umweltbundesamt.de/themen/verkehr-laerm/umgebungslaermrichtlinie/laermkarten*

Genaue Informationen gibt es auf den jeweiligen Seiten der Bundesländer. In Hessen zum Beispiel, werden im Download-Bereich EXCEL-Tabellen mit den jeweils aktuellen Belastetenzahlen für alle kartierten Gemeinden angeboten: *http://www.hlug.de/?id=525*

In Bayern werden beispielsweise Listen mit Betroffenenstatistiken verschiedener Kommunen veröffentlicht: *http://www.lfu.bayern.de/laerm/eg_umgebungslaermrichtlinie/kartierung/index.htm#betroffenheitsanalyse*

Treibhausgas

Um den Treibhausgasausstoß auszurechnen, gibt es Daten des Umweltbundesamtes für einzelne Verkehrsträger: *http://www.umweltbundesamt.de/themen/verkehr-laerm/emissionsdaten*

Vorschläge zur Berechnung von Treibhausgasemissionen von Spedition und Logistik liefert ein Leitfaden: *http://www.umweltbundesamt.de/sites/default/files/medien/461/publikationen/4306.pdf*

10.3 Kraftfahrzeugdaten

Das Kraftfahrt-Bundesamt (KBA) führt umfangreiche Statistiken zu Daten des Pkw- und Lkw-Bestands, den Antriebsarten, den Nutzlasten, zu Fahrerlaubnissen und Verkehrsauffälligkeiten und vieles mehr: *http://www.kba.de*

Der Verband der Automobilindustrie (VDA) stellt automobilspezifische Statistiken zusammen: *http://www.vda.de/*

10.4 Verkehrsdaten

Im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI) wird im Jahresrhythmus vom Deutschen Institut für Wirtschaftsforschung (DIW) Berlin eine umfassende verkehrsstatistische Datengrundlage im Taschenbuch »Verkehr in Zahlen« erarbeitet. Für nahezu alle Bereiche des Verkehrs in ihrer institutionellen und funktionalen Gliederung werden Statistiken für die Bundesrepublik Deutschland aufbereitet *(http://www.bmvi.de/viz)*.

Straßenverkehrszählungen, Straßendatenbanken, Unfallstatistiken

Die Bundesanstalt für Straßenwesen (BASt) stellt Statistiken zu Straßenverkehrszählungen und zu Verkehrs- und Unfalldaten zur Verfügung: *http://www.bast.de*

Weitere Verkehrsdaten liefern Straßenverkehrszählungen, Straßendatenbanken der Länder, amtliche Straßenverkehrsunfallstatistiken. Für die Recherche von Unfallschwerpunkten vor Ort kann Kontakt mit den örtlichen Polizeidienststellen aufgenommen werden.

Statistiken für den öffentlichen Personenverkehr

Statistiken des öffentlichen Personenverkehrs und des Schienengüterverkehrs aus Sicht der Verkehrsunternehmen stellt der Verband Deutscher Verkehrsunternehmen (VDV) vor, einem Zusammenschluss von ca. 600 lokalen Verkehrsunternehmen: *http://www.vdv.de*. Außerdem können viele Verkehrsverbünde in Deutschland sehr umfangreiche Statistiken liefern.

10.5 Verhaltensdaten

Mobilität in Deutschland (MiD)

Mobilität in Deutschland (MiD) ist eine bundesweite Befragung von 25.000 Haushalten zu ihrem alltäglichen Verkehrsverhalten im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI). Sie wurde erstmals im Jahr 2002 durchgeführt und im Jahr 2008 wiederholt. Ähnliche Umfragen fanden bereits 1976, 1982 und 1989 unter dem Namen »KONTIV« (Kontinuierliche Erhebung zum Verkehrsverhalten) statt. Die Daten bilden eine Grundlage für das Verkehrsverhalten in der Bundesrepublik Deutschland. Regionale Aufstockungen der Befragung liegen für den Raum Hannover und das Land Brandenburg vor. Die Daten können bei der Clearingstelle für Verkehrsdaten bestellt werden: *http://daten.clearingstelle-verkehr.de/order-form.html#223*

Die nächste Erhebung der MiD wird für das Jahr 2015 vorbereitet. Voraussichtlich werden die Ergebnisse für das Jahr 2016 erwartet.

Eine kostenfreie Webanwendung zur Auswertung der Erhebung Mobilität in Deutschland 2008 wird angeboten *(http://www.htmldb-hosting.com/pls/htmldb/f?p=MIT08:1)*.

Mobilität in Städten SrV 2008 (2013)

Die Verkehrserhebung Mobilität in Städten ist aus dem seit 1972 durchgeführten System repräsentativer Verkehrsverhaltensbefragungen (SRV) hervorgegangen und dient der Ermittlung von Mobilitätskennwerten der städtischen Wohnbevölkerung. Die Befragung wird vom Institut für Verkehrsplanung und Straßenverkehr der TU Dresden durchgeführt. Der Untersuchungsraum des SrV 2008 umfasste 76 Städte, Gemeinden und Verwaltungsgemeinschaften auf dem Gebiet der Bundesrepublik Deutschland. Davon befinden sich 55 in den neuen Bundesländern (einschließlich Berlin), 21 Städte und Gemeinden liegen im alten Bundesgebiet. Je nach Aufgabenstellung der Einzelstadt war eine vorgegebene Nettostichprobe zu erreichen, diese betrug für Oberzentren zwischen 1.000 Personen (Standardstadt) und 38.000 Personen (Berlin). In einigen Kommunen wurden die Daten in entsprechenden Publikationen oder Tabellenwerken aufbereitet. Inzwischen ist die Feldphase für die SrV 2013 abgeschlossen. Insgesamt wurde in über 100 Untersuchungsräumen der Bundesrepublik Deutschland eine Stichprobe von ca. 120.000 Personen befragt: *http://tu-dresden.de/die_tu_dresden/fakultaeten/vkw/ivs/srv*

Deutsches Mobilitätspanel

Das deutsche Mobilitätspanel wird vom Institut für Verkehrswesen am Karlsruher Institut für Technologie (KIT) wissenschaftlich betreut. Es handelt sich um ein rotierendes Panel, das seit 1994 durchgeführt wird. Hierbei werden die einzelnen Kohorten drei Jahre zu ihrem Verkehrsverhalten befragt. Die Grundgesamtheit bilden Haushalte in Deutschland. Die jährlichen Erhebungen haben einen Umfang von rund 1.500 Haushalten mit mehr als 3.000 Personen: *http://mobilitaetspanel.ifv.kit.edu/*

Kraftfahrzeugverkehr in Deutschland 2010

Die bundesweite Verkehrserhebung KiD 2010 ist Befragung von Kfz-Haltern über Einsatz und Nutzung von Kraftfahrzeugen. Die Befragung bietet Kennwerte wie beispielsweise Verkehrsaufkommen, Verkehrsmittelwahl und Fahrtenketten insbesondere beim Kraftfahrzeugverkehr bis einschließlich 3,5 t Nutzlast. Die Daten bieten einen grundlegenden Einblick in das Verkehrsverhalten von gewerblichen und privaten Pkw- und Lkw-Fahrerinnen und Fahrern. Die Daten können bei der Clearingstelle für Verkehrsdaten bestellt werden: *http://daten.clearingstelle-verkehr.de/order-form.html#223*

Fahrleistungserhebung 2002 (2014)

Die fahrzeugbezogene bundesweite Erhebung im Jahr 2014 richtete sich an circa 156.000 Fahrzeughalter. Berücksichtigt wurden alle Fahrzeugarten und Haltergruppen. Hier bildet die Fahrleistung von Kraftfahrzeugen (Kfz) auf dem Straßennetz innerhalb eines Jahres eine zentrale Kenngröße zur Beschreibung des motorisierten Straßenverkehrs. Die Erhebung wurde von der Bundesanstalt für Straßenwesen (BASt) durchgeführt. Fahrleistungserhebungen wurden für die Jahre 1990, 1993 und 2002 durchgeführt: *http://www.fahrleistungserhebungen.de/*

11 Abbildungen und Referenzen

Alle Cartoons stammen von AUKE HERREMA *(www.aukeherrema.nl)*.

Teile des Textes wurden 2013 im Buch »Evaluation Matters. A practitioner's guide to sound evaluation for urban mobility measures« im Waxmann Verlag veröffentlicht. Für diese deutsche umfassende Überarbeitung hat das Umweltbundesamt eine Lizenz von Waxmann erworben, welche die kostenlose Bereitstellung dieses Werkes erlaubt. Eine erste Übersetzung vom Englischen ins Deutsche erfolgte durch das Vertaalbureau Snelvertaler *(http://www.snelvertaler.nl/)*, eine weitere Übersetzung und Anpassung stammt von NICOLA MOCZEK.

12 Die Autorinnen und Autoren

KATRIN DZIEKAN

Promovierte Verkehrsplanerin und Diplom-Psychologin seit 2012 Fachgebietsleiterin Umwelt und Verkehr am Umweltbundesamt in Dessau-Roßlau. 2008–2012 Technische Universität Berlin, Fachgebiet Integrierte Verkehrsplanung u. a. Projektevaluationsmanagerin in CIVITAS MIMOSA. Schwerpunkte: Politikberatung, Verkehrsverhalten, nachhaltige Mobilität und Evaluation. *katrin.dziekan@uba.de*

VERONIQUE RIEDEL

Diplom-Ingenieurin, seit 2012 wissenschaftliche Mitarbeiterin am Fachgebiet Integrierte Verkehrsplanung der Technischen Universität Berlin. Schwerpunkte: Gamification, Mobilitätsbildung, quantitative Methoden, Präferenzmessungen. *veronique.riedel@tu-berlin.de*

STEPHAN DAUBITZ

Diplom-Politologe und Gesundheitswissenschaftler, seit 2008 wissenschaftlicher Mitarbeiter am Fachgebiet Integrierte Verkehrsplanung Technische Universität Berlin. Schwerpunkte: Mobilitätsarmut, empirische Mobilitätsforschung, qualitative und quantitative Methoden. *stephan.daubitz@tu-berlin.de*

NICOLA MOCZEK

Diplom-Psychologin, seit 1997 Geschäftsführerin von PSY:PLAN, Institut für Architektur- und Umweltpsychologie, Berlin. Schwerpunkte: Nachhaltige Stadtentwicklung, Wohnen, Mobilität, Nutzung erneuerbarer Energien, Orientierungssysteme, Besucherforschung und Evaluation. Seit 2001 Mitherausgeberin der Zeitschrift Umweltpsychologie. *moczek@psyplan.de*

STEPHANIE KESSLER

Diplom-Geographin, seit 2011 freiberufliche Beraterin für nachhaltige städtische Mobilität. 2009–2013 Mitarbeit in CIVITAS POINTER (evaluation support). Schwerpunkte: nachhaltige, städtische Mobilität im internationalen Vergleich, EU-Förderung, qualitative Evaluierung, Kampagnen, Fahrradverkehr. *mail@stephaniekessler.eu*

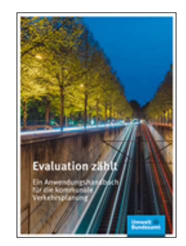
STEFANIE KETTNER

Freiberuflich tätige Geographin im Bereich Verhaltens- und Marktforschung mit den Schwerpunkten Mobilität und Tourismus. 2009–2014 wissenschaftliche Mitarbeiterin am Fachgebiet Integrierte Verkehrsplanung, Technische Universität Berlin. *stefanie.kettner@gmx.de*

MICHAEL ABRAHAM

Diplom-Ingenieur Stadt- und Regionalplanung, seit 2012 wissenschaftlicher Mitarbeiter am Zentrum für Technik und Gesellschaft (ZTG) der Technischen Universität Berlin, im Bereich Mobilität und Raum. 2007–2012 Technische Universität Berlin, Fachgebiet Integrierte Verkehrsplanung u. a. Mitarbeit in CIVITAS MIMOSA. Schwerpunkte: Prozessevaluation, nachhaltige Mobilität, Fahrradverkehr, Verkehrsplanung im internationalen Bereich. *info@michaelabraham.de*





► Die Broschüre als Download
[Kurzlink: http://www.uba.de/evaluation-verkehrsplanung]